ESTADISTICA ESPAÑOLA
Vol. 32, Núm. 124, 1990, págs. 327 a 388

Cointegración: Una panorámica (*)

JUAN J. DOLADO
Servicio de Estudios
Banco de España

Marzo 1989

RESUMEN

Este trabajo contiene una panorámica de desarrollos recientes en la teoría de la cointegración entre un conjunto de series temporales integradas. La cointegración de dichas series implica la existencia de relaciones entre las mismas que tiene carácter estacionario y por tanto representa una forma natural de definir relaciones de equilibrio a largo plazo entre las variables. En el trabajo se recogen aportaciones recientes referentes a la representación, estimación, contrastación y predicción en sistemas cointegrados, incorporando la posibilidad de cointegración en diferentes frecuencias temporales.

Clasificación AMS: 90A20, 62M20, 62P20.

Palabras clave: Cointegración, Tendencias, Mecanismo de Corrección del Error.

(*) Este artículo ha sido preparado para *Estadística Española*. Su contenido se ha visto beneficiado por comentarios de A. Banerjee, A. Escribano, J. Galbraith, D. Hendry, T. Jenkinson, D. Peña y participantes en seminarios en las universidades de Londres, París y Valencia, en el Banco de España y en el Instituto Nacional de Estadística.

1. INTRODUCCION

La mayoría del trabajo econométrico está asentado sobre el supuesto de estacionariedad. Hasta hace poco tiempo este supuesto apenas se cuestionaba y el análisis económico procedía como si todas las series temporales económicas fuesen estacionarias, al menos alrededor de tendencias determinísticas. Las series estacionarias deben tener, al menos, medias y varianzas incondicionales constantes, un supuesto que raramente se verifica en economía. La importancia de dicho supuesto ha sido reconocida durante mucho tiempo, sobre todo desde que los importantes artículos de Granger y Newbold (1974) y Nelson y Plosser (1982) pusieron en alerta a la profesión sobre las consecuencias de la no estacionariedad. Las variables *integradas* constituyen una clase específica de variables no estacionarias con importantes propiedades tanto desde el punto de vista estadístico como económico. Estas propiedades se derivan de la presencia de tendencias estocásticas (raíces unitarias), en vez de únicamente determinísticas, teniendo las innovaciones de un proceso integrado carácter permanente en vez de transitorio. Por ejemplo, en términos de análisis del bienestar, ello implica que los costes de errores de expectativas producidos, por ejemplo, por variaciones en la política económica tienen un carácter mucho más relevante que en el caso en que las perturbaciones fuesen puramente transitorias.

En particular, la presencia de raíces unitarias en muchas series económicas resulta ser una implicación teórica del uso racional de la información disponible por parte de los agentes económicos. Ejemplos típicos entre variables financieras son los contratos a futuros, los precios de las acciones, los tipos de interés nominales y reales y el tipo de cambio, mientras que entre las variables reales encontramos el paro (histéresis), la inversión y el consumo. Quizás sea este último el ejemplo más conocido, derivado de añadir expectativas racionales a la teoría de la renta permanente. A la vista de esta epidemia de raíces unitarias en economía, ha aparecido una literatura voluminosa en estos últimos años sobre la contrastación, estimación y especificación de modelos que contienen variables integradas, por lo que el propósito de esta panorámica es proporcionar una guía a través de esta literatura crecientemente compleja. Con el fin de limitar en lo posible el contenido de este artículo, nos circunscribiremos al análisis de la cointegración, es decir, al análisis de modelos multivariantes con series integradas.

Resulta comprensible que en estos últimos años se haya producido un importante progreso asociado con el tema de la cointegración [1]. La razón principal de este interés creciente reside en el hecho de que esta rama de investigación abre la posibilidad de estimar y contrastar la existencia de relaciones de equilibrio a largo plazo entre diversas variables, tal como sugiere la teoría económica. Tal como comentábamos previamente, la evidencia empírica sugiere que la mayoría de las series económicas son integradas, requiriendo un cierto grado de diferenciación a fin de convertirlas en estacionarias. Sin embargo, la teoría económica raramente sugiere relaciones de equilibrio que no sean funciones estacionarias de las variables originales, esto es, sin transformar. Ello implica que, de acuerdo con la teoría económica, existen mecanismos económicos que hacen que, a lo largo del tiempo, las variables se comporten estocásticamente del mismo modo. En otras palabras, mientras que las variables que conforman el modelo pueden ser todas ellas integradas, las desviaciones respecto a determinadas relaciones de equilibrio, sugeridas por la teoría resultan ser estacionarias.

Se reconoce desde hace mucho tiempo que las técnicas usuales de análisis de regresión, en presencia de variables integradas, puede llevarnos a conclusiones erróneas, incluso cuando el análisis formal de este hecho sea reciente (véase Phillips (1986))[2]. El problema de las regresiones "espúreas", tal como fue puesto de manifiesto por Granger y Newbold (1974), produjo un grado creciente de desconfianza en el análisis econométrico. Ello condujo a muchos economistas a adoptar versiones extremas del enfoque de Box y Jenkins (1976) a través del cual todas las series se convertían en series estacionarias, previamente al análisis de regresión, con lo que, en la mayoría de los casos, sólo se consideraban variables diferenciadas (por ejemplo, en el enfoque de "funciones de transferencia"). Ello implicaba que, en general, este tipo de metodología ignoraba las propiedades a largo plazo del modelo ignorando, por tanto, las relaciones de equilibrio sugeridas por la teoría económica. Todo ello dificultaba, en muchos casos, la interpretación de los modelos.

Una importante línea de investigación que intentaba ofrecer una respuesta a dichos problemas, se basó en el uso de modelos de *mecanismo de corrección del error* (ECM) en la modelización econométrica. La idea de modelar ecuaciones de comportamiento económico en forma de "servomecanismos" que traten de corregir de forma sistemática errores cometidos en el pasado, se remonta al trabajo de Sargan (1964), basado en ideas de Phillips (1954), y tiene la ventaja de retener información sobre el nivel de las variables y, consecuentemente, sobre las relaciones de equilibrio entre dichas variables. Numerosos ejemplos de aplicaciones satisfactorias de este tipo de metodología han aparecido posteriormente en la literatura (véase, por ejemplo, inter alia. Davidson et al. (1978), Hendry y Mizon (1978), Davidson y Hendry (1981) Currie (1981) y Salmon (1982)), mientras que el análisis estadístico subyacente ha sido desarrollado por Hendry y Richard (1983, 1984). Aún así, existía una laguna en la comprensión de cómo los modelos de regresión con variables integradas podían conducir a términos de error estacionarios e incluso ir más allá y suponer perturbaciones i.i.d. Esta laguna ha sido cubierta mediante un importante artículo de Granger (1981), donde se establece, aunque de manera informal, la equivalencia entre los conceptos de cointegración y modelos ECM. Posteriormente, Granger y Weiss (1983) y Engle y Granger (1987) han desarrollado formalmente las ideas anteriores a través de un conjunto de teoremas de representación que establecen las conexiones existentes entre dichos conceptos. En resumen, se demuestra que un modelo ECM produce un conjunto de variables cointegradas y que la existencia de dicho conjunto implica una representación de las variables en forma de un modelo ECM. En gran medida, el concepto de cointegración provee un soporte estadístico formal al uso de modelos ECM y proporciona adicionalmente un procedimiento genérico para contrastar la validez de las predicciones, en términos de relaciones de equilibrio, derivadas de diversas teorías económicas.

El artículo está organizado del siguiente modo. La Sección 2 ilustra brevemente el marco estadístico necesario para analizar series integradas, concentrándose, en aras de simplicidad, en el caso más simple de integración, i.e. el paseo aleatorio. La Sección 3 analiza las restricciones que la existencia de cointegración impone sobre las diversas representaciones de un sistema multivariante, así como los métodos de descomposición de dicho sistema en sus varios componentes integrados. Se presta especial atención a sistemas integrados de primer orden, aunque otros casos más generales también son brevemente analizados. La Sección 4 trata de estimación, contrastación y predicción en sistemas cointegrados. Finalmente, la sección 5 extiende parte del análisis previo a un marco de máxima verosimilitud, un enfoque desarrollado muy recientemente y que parece ser una de las líneas de investigación más fructíferas para el futuro.

2. PROPIEDADES ESTADISTICAS DE VARIABLES INTEGRADAS

Una serie temporal debilmente estacionaria debería tener media y varianza incondicionales invariantes con el tiempo. Sin embargo, la mayoría de las series económicas no satisfacen esta propiedad, al tener primer y segundos momentos que aumentar con el tamaño muestral (véase Escribano (1987) para la definición precisa de integración en el momento *i-smo* de un proceso estocástico)[3]. Tales series son no-estacionarias y pueden requerir un cierto grado de diferenciación para lograr estacionariedad. Una variable que requiere d diferencias para inducir estacionariedad se denomina *integrada de orden d* y se denota como I(d) (véase Granger (1982)). Un ejemplo de serie I(1) es el paseo aleatorio con deriva, cuyo *proceso generador de los datos* (PGD) es

$$\Delta y_t = -\alpha y_{t-1} + \mu + \varepsilon_t; \ \alpha = 0, \ y_0 = 0, \ \varepsilon_t \sim niid(0, \sigma_\varepsilon^2) \tag{1}$$

o

$$y_t = \mu t + S_t; \quad S_t = \sum_1^t \varepsilon_i \tag{2}$$

Nos concentraremos en esta sección en las propiedades estadísticas que se derivan de la presencia de una única raíz unitaria como en (1), y para simplificar supondremos que μ=0. En general, las series integradas, como y_t, resultan ser funciones lineales de una tendencia (si $\mu \neq 0$). Las desviaciones de dicha tendencia, S_t, son no-estacionarias al resultar de la acumulación de perturbaciones aleatorias pasadas, dando lugar al concepto de series integradas.

Con el fin de generalizar la especificación del PGD, abandonamos el supuesto simplificador de perturbaciones *iid* y pasamos a imponer ciertas condiciones más generales sobre la secuencia $\{\varepsilon_t\}$. Estas condiciones son necesarias si se quiere obtener distribuciones no degeneradas para el conjunto de estadísticos descritas posteriormente. El conjunto más débil de condiciones que puede imponerse se encuentra definido detalladamente en Phillips (1987) y puede resumirse del siguiente modo

(a) $E(\varepsilon_t) = 0 \ \forall t$

(b) $sup\ E|\varepsilon_t|^\beta < \infty$ para algún $\beta > 2$

(c) $\sigma^2 = lim\ E(T^{-1} S_T^2)$ existe y $\sigma^2 > 0$

(d) ε_t tiene una mixtura fuerte con coeficientes de mixtura α_m, tales que $\Sigma\, \alpha_m^{(1-2/\beta)} < \infty$.

La condición (b) restringe el grado de heterogeneidad del proceso, mientras que (c) controla la normalización en un grado que permita la existencia de distribuciones asintóticas no degeneradas. La varianza asintótica σ^2 es igual a $2\pi f(0)$, siendo $f(0)$ el espectro de ε_t evaluado en la frecuencia cero. La condición (d) modera el grado de dependencia temporal en relación con la probabilidad de observaciones atípicas (véase White (1984)).

La generalidad del anterior conjunto de condiciones implica que el modelo (1) engloba una gran variedad de PGD's. Entre estos se encuentra virtualmente cualquier modelo ARIMA con una raíz unitaria e incluso modelos ARMAX con una raíz unitaria y procesos exógenos no integrados.

Con el fin de derivar las distribuciones asintóticas que nos interesan, es necesario, tal como ocurre en el marco estacionario, utilizar una sucesión de variables aleatorias cuya convergencia se asegura a través de determinadas transformaciones. En términos más precisos, en el marco no estacionario necesitamos construir la sucesión de sumas parciales $\{S_t\}$ que ha de ser transformada de manera que sus elementos pertenezcan al espacio D(0,1) de todas las funciones de valor real en el intervalo [0,1] que son contínuas por la derecha y tienen límites finitos por la izquierda. Esto se logra mediante la sustitución de la serie original $\{S_t\}$ por la sucesión concentrada

$$S_T(r) = T^{-1/2}\, S_{[Tr]},\ r\varepsilon[0,1] \tag{3}$$

donde $[z]$ representa la parte entera de z. De esta manera se concentra el dominio de $\{y_t\}$ en el intervalo [0,1], indiciando las observaciones mediante r. As, por ejemplo, si T=100, la observación original y_{50} estará indiciada por $r\varepsilon(0{,}50,\ 0{,}51)$. La elección de $T^{-1/2}$ como grado de normalización asegura que la serie transformada no sea ni explosiva ni converja a cero. Por ejemplo si ε_t es *iid*, entonces $var(y_T) = T\sigma^2_\varepsilon$, con lo que su desviación típica será de orden $Op(T^{1/2})$ y, por tanto, este será el orden elegido para modificar el rango de $\{y_t\}$.

Bajo el conjunto de condiciones (a)-(d), tendremos que cuando $T \Rightarrow \infty$, $S_T(r) \Rightarrow \sigma W(r)$, donde "$\Rightarrow$" denota convergencia débil en probabilidad. $W(r)$ es un proceso de Wiener (o movimiento Browniano) con varianza unitaria, que pertenece al espacio C[0,1] de funciones de variable real en el intervalo [0,1]. Este resultado de convergencia se conoce como Teorema de Donsker (véase Billingsley (1968)). Nótese que $W(r)$ se comporta como un paseo aleatorio en tiempo contínuo, tal que para r fijo se distribuye como $N(0, r)$ y tiene incrementos independientes. Además, una extensión del teorema de Slutsky para variables aleatorias estacionarias, también se verifica en este contexto, en el sentido de que si $g(.)$ es cualquier función contínua en C[0,1], entonces $S_T(r) \Rightarrow \sigma W(r)$ implica que $g[S_T(r)] \Rightarrow g[\sigma W(r)]$. Este resultado se conoce como Teorema de la Correspondencia Continua (véase Billingsley (1968)). La diferencia más importante entre la teoría asintótica aplicable a procesos estacionarios y la aplicable a procesos integrados, reside en el hecho de que mientras en la primera los momentos muestrales convergen a constantes, en la segunda convergen a variables aleatorias. Igualmente, como resultado de la ausencia de estacionariedad y ergodicidad, los Teoremas Centrales del límite tradicionales se sustituyen por Teoremas Funcionales del Límite (véase Billingsley (1968))[4].

Como ilustración de los resultados anteriores, los siguientes momentos muestrales estandarizados convergen a funcionales de Wiener

(i) $T^{-2} \Sigma\, y_t^2 \Rightarrow \sigma^2 \int_0^1 W(r)^2\, dr$

(ii) $T^{-3/2} \Sigma y_t \Rightarrow \sigma \int_0^1 W(r)\, dr$

(iii) $T^{-1} \Sigma y_{t-1} \varepsilon_t \Rightarrow \sigma^2 / 2 [W(1)^2 - \sigma_\varepsilon^2 / \sigma^2]$

Nótense las divergencias entre las órdenes de magnitud de estos procesos límite con los correspondientes a los procesos estacionarios convencionales, *i.e.* $Op(T^2)$ en vez de $Op(T)$ en (i), $Op(T^{3/2})$en vez de $Op(T)$ en (ii) y $Op(T)$ en vez de $Op(T^{1/2})$ en (iii). Estas diferencias iluminan las características no convencionales de la consistencia y distribución de los coeficientes que sirven como base para contrastar raíces unitarias y que serán de gran importancia en la discusión posterior del concepto de cointegración.

Si, por ejemplo, se aplican mínimos cuadrados ordinarios (MCO) a (1), resulta fácil derivar en base a los procesos límites (i)-(iii), que el estimador de la pendiente, $\hat{\alpha}$, y su *t*-ratio convergen a las siguientes distribuciones límite (cuando μ=0)

$$T \hat{\alpha} \Rightarrow \frac{1/2\, [W(1)^2 - \sigma_\varepsilon^2 / \sigma^2]}{\int_0^1 W(r)^2\, dr} \tag{4}$$

$$t_{\alpha=0} \Rightarrow \frac{1/2\, \sigma\, [W(1)^2 - \sigma_\varepsilon^2 / \sigma^2]}{\sigma_\varepsilon\, [\int_0^1 W(r)^2\, dr]^{1/2}} \tag{5}$$

En (4) se observa que $\hat{\alpha}$ converge a su verdadero valor, α=0, a velocidad $Op(T^{-1})$ en vez del orden convencional $Op(T^{-1/2})$. Igualmente, en (5), se observa que el *t*-ratio correspondiente tiene una distribución límite no degenerada que difiere de la distribución normal estandarizada que se utiliza en teoría asintótica convencional. Las distribuciones empíricas de (4) y (5) han sido obtenidas por simulación y constituyen la base de los contrastes de raíces unitarias, tanto en sus versiones paramétricas (véase, inter alia, Fuller (1976) Dickey y Fuller (1979, 1961), Evans y Savin (1981), Dickey et al. (1986) y Dickey y Pantula (1987)), como en las no paramétricas (véase Phillips y Ouliaris (1988) y Perron (1988)). Tal como se hizo notar en la Introducción este es un tema que no se trata en la presente panorámica.

3. COINTEGRACION Y SUS DIFERENTES REPRESENTACIONES

En esta sección se extiende el concepto de proceso integrado, examinado en la sección anterior, a un contexto multivariante, analizando el concepto de *cointegración*, es decir la posibilidad de la existencia de relaciones lineales entre variables integradas que resultan tener un orden menor de integración. El caso más interesante es aquel en que la relación existente es estacionaria, I(O), ya que en este caso puede aplicarse el análisis de regresión convencional. Además, se discuten las diferentes representaciones de procesos cointegrados, entre los cuales la más útil es aquella en términos de modelos ECM, donde los agentes económicos utilizan "servomecanismos" correctores de los errores cometidos previamente, de acuerdo con una relación de equilibrio preestablecido.

3.1. Descomposición de una serie Temporal en Componentes Integrados

Comenzamos por extender la definición de proceso integrado a un vector de series temporales. Sea $\{x_t\}$ un proceso n-dimensional tal que, en ausencia de componentes determinísticos, tiene la siguiente representación ARIMA multivariante

$$\Phi(L)\, x_t = \theta(L)\varepsilon_t;\ \varepsilon_t \sim NI(0,\Omega) \tag{6}$$

donde $\Phi(L)$ y $\theta(L)$ son $(n \times n)$ matrices de polinomios en el operador de desfases L, de orden p y q respectivamente. El polinomio autorregresivo es tal que *det* $\Phi(L)$ tiene todas sus raíces fuera del círculo unitario, excepto un número d de raíces iguales a la unidad y $\Phi(0) = \theta(0) = I$

De acuerdo con (6) definiremos a $\{x_t\}$ como un proceso integrado de orden d, $I(d)$, si tiene la siguiente representación de Wold

$$(1-L)^d x_t = C(L)\ \varepsilon_t \tag{7}$$

donde $C(L) = \Phi^*(L)\ \theta(L)/[\det \Phi(L)/(1-L)^d]$, siendo $\Phi^*(L)$ la matriz adjunta de $\Phi(L)$. C(L) es una matriz de fracciones racionales que admiten un desarrollo en serie con coeficiente decrecientes en forma exponencial. Nótese que el álgebra de los procesos integrados es tal que si d_1 y d_2 representan los respectivos órdenes de integración de dos series temporales, entonces

$$I(d_1) + I(d_2) = I\{d \leq max\ (d_1, d_2)\} \tag{8}$$

A partir de la representación de Wold (7), el proceso $\{x_t\}$ puede descomponerse en sus componentes integrados de orden decreciente de la siguiente forma

$$x_t = C(1)(1-L)^{-d}\varepsilon_t - C^{(1)}(1)(1-L)^{-(d-1)}\varepsilon_t + \ldots + \frac{C^{d-1}(1)(-1)^{d-1}}{(d-1)!}\ (1-L)^{-1}\varepsilon_t + C^*(L)\varepsilon_t \tag{9}$$

Esta representación se obtiene mediante un desarrollo de Taylor de la matriz C(L) hasta el orden (d-1), dividiendo posteriormente por $(1-L)^d$. Resulta fácilmente demostrable que la matriz residual C*(L) tiene un determinante cuyas raíces se encuentran fuera del círculo unitario. De esta manera se logra una descomposición de x_t en términos de sus varios componentes integrados, que puede reescribirse como

$$x_t = x_t^{(d)} + x_t^{(d-1)} + \ldots + x_t^{(1)} + x_t^{(0)} \tag{10}$$

donde $x_t^{(k)} = [(-1)^{d-k}\, C^{(d-k)}(1)\, (1-L)^{-k} / (d-k)!\,]\, \varepsilon_t$, $(k > 0)$

Resulta evidente de (10), que la descomposición no es única, ya que siempre pueden reordenarse los diferentes componentes, por ejemplo $(x_t^{(d-k)} + x_t^{(d-k-1)} / 2)$ de forma que tenga el mismo orden de integración que $x_t^{(d-k)}$. Entre las infinitas posibles descomposiciones de x_t, se ha escogido aquella en que cada uno de los componentes tiene la representación ARIMA más sencilla. Resulta claro que (10) verifica esta propiedad y, en este sentido, se denominará *descomposición canónica.*

Seguidamente analizamos la existencia de combinaciones lineales de los elementos de x_t, tales que algunos (o todos) de los componentes integrados desaparezcan. En este sentido, se define a x_t como ***cointegrada*** de orden k, C(d,k) si existe un vector ($1 \times n$), α, tal que $\alpha' x_t$ sea I(d-k). Esto es, a partir de (10)

$$\alpha' x_t = \alpha' x_t^{(d-k)} + \ldots + \alpha' x_t^{(0)} \sim I(d-k) \tag{11}$$

La descomposición (11), de acuerdo con (10), implica que

$$\alpha' C(1) = \alpha' C^{(1)}(1) = \ldots = \alpha' C^{(d-k+1)}(1) = 0 \tag{12}$$

es decir $ker\ C^{(i)}(1)' \neq \{0\}$ (i=1,..., d-k+1). Por tanto si definimos los espacios vectoriales

$$C_s = \bigcap_{k=0}^{s} ker\,[C^{(k)}(1)'] \tag{13}$$

podemos alternativamente definir a x_t como cointegrada C(d,k) si y solo si existe un vector α tal que $\alpha \in C_{d-k+1}$ y $\alpha \notin C_{d-k}$

Ejemplo 1: Si x_t es C(1,1), ello implica que existe, al menos, un vector α tal que $\alpha' C(1)=0$.

La relación existente entre los conceptos de cointegración y equilibrio a largo plazo aparece ahora más nítida. Una manera natural de definir el equilibrio entre un conjunto de variables es postular una relación del tipo

$\alpha' x_t=0$. Por ejemplo, si se supone que una cierta proporción γ de los aumentos en la productividad laboral (q en logaritmos) se traslada a aumentos en el salario real ($(w$-$p)$ en logaritmos), entonces, en equilibrio (aparte de constantes), se verifica la relación w-p=γq. Por tanto si w-p-γq=0 en cualquier período t, el mercado de trabajo estará en equilibrio (nótese que si γ=1, la relación previa representa la condición de primer orden de una empresa competitiva con una función de producción Cobb-Douglas). Es claro que los salarios reales tenderán a responder con cierto retraso a los cambios en la productividad y la forma en que se restaura el equilibrio puede ser un proceso complejo. En este caso, la perturbación $e_t=w_t-p_t-\gamma q_t$ mide la desviación de la relación de equilibrio, o desequilibrio, en el período t. Si $(w\text{-}p)_t$ y q_t están cointegradas, entonces, de acuerdo con las definiciones anteriores, la desviación e_t ha de ser I(0). Una forma obvia de contrastar esta teoría es determinar el orden de integración de e_t. Si no es posible rechazar la hipótesis nula de una raíz unitaria en e_t, no existirá un mecanismo de ajuste de los salarios reales a su valor de equilibrio, en cuyo caso la relación postulada constituye un concepto irrelevante. La teoría económica sugiere muchas relaciones de equilibrio similares a la anterior, por ejemplo, consumo y renta disponible (Campbell (1987) y Banerjee y Dolado (1988)); tipo de cambio nominal y nivel de precios relativos (Edison y Kloveland (1987)); renta nominal, tipo de interés y saldos nominales de dinero (Hendry y Ericsson (1986) y Dolado (1988)), tipos de interés a corto y largo plazo (Campbell y Shiller (1987)), etc. La contrapartida estadística a las anteriores relaciones de equilibrio es la existencia de cointegración entre las variables previamente mencionadas.

El ejemplo anterior tiene también la virtud de ilustrar varias características del concepto de cointegración. Así, por ejemplo, si w_t y p_t están cointegradas C(d,b), el vector de cointegración ha de ser único ya que si existiese cualquier otra combinación lineal independiente de la anterior, siempre se podría expresar una variable en términos de la otra añadiendo o sustrayendo w_t(o p_t) que es I(d). Sin embargo, cuando se trata de más de dos variables, por ejemplo w_t, p_t y q_t, entonces puede existir más de un vector de cointegración. Así $p_t-w_t=-\beta q_t$ representaría una ecuación de precios de "mark-up", que coexiste con la relación de salarios. Nótese que si existiesen, en este caso, tres vectores independientes, entonces las tres variables serían estacionarias y cualquier relación entre ellas también sería estacionaria, con lo que el concepto de cointegración está vacío de contenido.

La existencia de cointegración también posibilita la identificación de modelos que en un contexto estacionario no estarían identificados. Para ilustrar este hecho, considérese el siguiente ejemplo tomado de Engle y

Granger (1987). Supóngase que $x_t=(y_t, z_t)'$ está generado de acuerdo con el siguiente PGD

Ejemplo 2

$$\begin{aligned} & y_t + \beta z_t = v_t \;\; ; \;\; v_t = \rho_1 v_{t-1} + \varepsilon_{1t} \\ & y_t + \alpha z_t = u_t \;\; ; \;\; u_t = \rho_2 u_{t-1} + \varepsilon_{2t} \\ & \varepsilon_{it} \sim niid(0,\sigma_i^2) \; (i=1,2) \;\; ; \;\; E(\varepsilon_{it}\, \varepsilon_{js}) = 0 \end{aligned} \tag{14}$$

Resulta claro que si y_t y z_t, fuesen estacionarios, entonces α y β no estarían identificados. Sin embargo, dependiendo del valor tomado por los parámetros $\rho_i's$, surgen los siguientes cuatro casos

(i) $\rho_1=1$, $\rho_2 < 1$; implica que y_t y z_t son I(1) y el vector de cointegración es $(1,\alpha)$. Nótese que para un tamaño muestral T suficientemente grande, la perturbación v_t tendrá una varianza que tiende a infinito. Por tanto, cualquier método de minimice la suma de cuadrados de los residuos, ponderada o no ponderada, identificará la segunda ecuación con una muestra suficientemente grande.

(ii) $\rho_1 < 1$, $\rho_2 = 1$; implica, simétricamente al caso anterior, que y_t y z_t son I(1) y el vector identificado de cointegración es $(1,\beta)$

(iii) $\rho_1=\rho_2=1$; implica que y_t y z_t son I(1) pero no existe ningún vector de cointegración

(iv) $\rho_1 < 1$, $\rho_2 < 1$; implica que y_t y z_t son I(0), lo que provoca falta de identificación.

Finalmente, nótese que aunque la discusión previa ha sido desarrollada en términos de un marco de relaciones estáticas entre variables integradas, no existe razón alguna para ignorar la existencia de relaciones dinámicas de cointegración. Por ejemplo, supóngase el siguiente PGD

Ejemplo 3

$$\begin{aligned} & y_t = \beta z_t + \gamma \Delta z_t + \varepsilon_{1t} \\ & \Delta^2 z_t = \varepsilon_{2t} \\ & \varepsilon_{it} \sim niid\,(0,\sigma_i^2) \; (i=1,2), \; E(\varepsilon_{it}\, \varepsilon_{js}) = 0 \end{aligned} \tag{15}$$

Se deriva de (15) que y_t y z_t son I(2) y de la primera ecuación se deriva el vector de cointegración $(1,-\beta,-\gamma)$ que produce una combinación lineal de y_t, z_t y Δz_t que es I(0). Nótese que y_t y z_t son C(2,1) y, por tanto, necesitan asociarse con otra variable I(1), Δz_t en este caso, para lograr una condición necesaria de cointegración.

3.2. Teoremas de Representación

Una vez que el concepto de cointegración ha sido analizado, es importante examinar las restricciones que impone sobre las diferentes representaciones AR o MA que se derivan de la representación general ARIMA expresada en (6). En línea con el artículo de Engle y Granger (1987), comenzaremos por analizar el caso más sencillo de la clase de Teoremas de Representación, que corresponde a que x_t sea C(1,1). Por tanto cada elemento del vector x_t, x_{it}, es I(1) y existen combinaciones lineales de dichos elementos que son I(0). El siguiente conjunto de resultados se conoce en la literatura como Teorema de Representación de Granger [5].

Sea $\{x_t\}$ un proceso estocástico n-variante C(1,1), tal que existen r vectores de cointegración ($r \leq n$-1) agrupados en la matriz (nxr) Γ. Entonces existen las siguientes representaciones

i) Representación MA

A partir de (7) con d=1, x_t sigue el proceso MA

$$(1-L)x_t = C(L)\varepsilon_t \tag{16}$$

y de acuerdo con (11)

$$\alpha' C(1) = 0 \Rightarrow \Gamma' C(1) = 0 \tag{17}$$

Por tanto, las filas de Γ pertenecen al espacio *ker* de C(1)'. Como el número de filas en C(1) es n, igual al rango más la dimensión del espacio *ker* (véase, por ejemplo Dhrymes (1978)), se sigue que

$$rango\ C(1) = n-r \tag{18}$$

por lo que C(1) no tiene rango completo

ii) Representación AR

La representación (16) puede reescribirse en forma AR

$$A(L)x_t = \varepsilon_t \tag{19}$$

donde, siendo I la matriz identidad,

$$A(L)\ C(L) = (1-L)I \Rightarrow A(1)\ C(1) = 0 \tag{20}$$

Por tanto, las filas de A(1) pertenecen al espacio *ker* de C(1)' y a partir de (17) puede escribirse del siguiente modo:

$$A(1) = B\Gamma' \tag{21}$$

donde B es una matriz ($n \times r$). Puesto que A(1) ε *ker* C(1)' se sigue que

$$rango\ A(1) = r \qquad (22)$$

Dado que A(1) es singular, la representación VAR será tal que los procedimientos numéricos utilizados para calcular la representación MA resultarán altamente inestables, signo de multicolinealidad, lo que dificulta el análisis de multiplicadores en las funciones impulso-respuesta.

iii) Representación ECM

Utilizando un desarrollo de primer orden de Taylor de A(L) alrededor de L=1, se obtiene

$$A(L) = A(1) + A^*(L)\ (1\text{-}L) \qquad (23)$$

donde el determinante de A*(L) tiene sus raíces fuera del círculo unitario. Sumando y restando A(1) L a (23) y utilizando (21), se obtiene

$$A(L)x_t = [B\Gamma'\ L + A^{**}(L)\ (1\text{-}L)]\ x_t = \varepsilon_t \qquad (24)$$

donde $A^{**}(L) = A(1)+A^*(L)$. Por tanto $A^{**}(0) = A(1) + A^*(0)$ y, dado (23), $A^{**}(0) = I$. Consecuentemente, (24) puede reescribirse en términos de un modelo multivariante ECM

$$A^{**}(L)\ \Delta\ x_t = -B\Gamma'\ x_{t-1} + \varepsilon_t \qquad (25)$$

Nótese que en la representación (25) la existencia de cointegración se deriva de la presencia de los niveles de las variables, de forma que la estimación de modelos VAR en diferencias dará lugar a errores de especificación si las variables están cointegradas. Igualmente la estimación de modelos VAR en niveles que ignoren las restricciones (21) presentan graves problemas al obtener la representación MA, tal como se comentó previamente. Sin embargo, Sims et al (1986) han demostrado recientemente que en un modelo VAR donde existen relaciones de cointegración entre las variables, siempre que cualquier variable pueda ser escrita como variable estacionaria con media cero, por ejemplo Δy_t, Δx_t y $(y_{t-1}-\alpha x_{t-1})$ y sus desfases, en un modelo bivariante, sus coeficientes se distribuirán normalmente de acuerdo con la teoría asintótica convencional. Como veremos posteriormente, esto no es completamente cierto cuando se trate de los coeficientes del propio vector de cointegración.

La representación (25) también ilustrá la conexión existente entre los conceptos de cointegración y "causalidad de Granger". Por ejemplo, si se trata de un sistema bivariante deberá haber causalidad en la dirección $z \rightarrow y$ si el segundo elemento de la matriz B, que es (2×1), es nulo, o en la dirección $y \rightarrow z$ si es el primer elemento el que es nulo. Si ambos elementos de B difieren de cero, entonces existe causalidad en ambas direcciones y ninguna de las variables puede considerarse como débilmente exógenea con respecto a los parámetros de la otra ecuación (véase Engle et al. (1983)). En este último caso existen restricciones no lineales a través de las ecuaciones, y para su estimación eficiente se necesita un procedimiento de información completa

iv) Representación TC (Tendencias Comunes)

La idea subyacente a esta representación, tal como ha sido sugerida por Aoki (1987) y Stock y Watson (1989), es que si un grupo de variables están cointegradas, deberá existir un conjunto de tendencias estocásticas comunes que guíen las representaciones ARIMA de las variables individualmente. Estas tendencias comunes desaparecerán mediante el uso de determinadas combinaciones lineales entre las variables. Con el fin de ilustrar este método de descomposición, utilicemos una versión de primer orden de (10) cuando d=1, obteniendo

$$x_t = C(1)\, x_t^{(1)} + C^*(L)\, \varepsilon_t = C(1) x_t^{(1)} + \sum_1^t [\, \sum_0^{t-i} C_j - C(1)]\, \varepsilon_i \qquad (26)$$

o bien

$$x_t = C(1)\, x_t^{(1)} - \sum_2^t [\sum_{t-i+1}^{t-1} C_j]\, \varepsilon_i \qquad (27)$$

Las tendencias estocásticas vienen dadas por $x_t^{(1)} = \sum_1^t \varepsilon_i$, mientras que el segundo término en (27) es estacionario, para t suficientemente grande, dado que las matrices C_j decaen exponencialmente para j suficientemente grande. A partir de (18) sabemos que el rango de C(1) es n-r y por tanto existen (n-r) tendencias comunes en el sistema, de ahí la denominación TC para esta representación.

Con el fin de ilustrar las diferentes representaciones examinadas en esta sección, considérese el siguiente ejemplo

Ejemplo 4

Supongamos que $x_t = (y_t, z_t)$ tiene la siguiente representación MA

$$(1-L) \begin{pmatrix} y_t \\ z_t \end{pmatrix} = (1-0{,}2L)^{-1} \begin{pmatrix} 1-0{,}6L & 0{,}8L \\ 0{,}2L & 1-0{,}6L \end{pmatrix} \varepsilon_t$$

Así pues

$$C(1) = \begin{pmatrix} 0{,}5 & 1{,}0 \\ 0{,}25 & 0{,}5 \end{pmatrix} \rightarrow \text{rango } C(1) = 1$$

Por tanto, y_t y z_t son C(1,1). El vector de cointegración, normalizando el coeficiente de y_t, es $\alpha' = (1,-2)$

Invirtiendo la matriz MA, se obtiene la representación AR

$$\begin{pmatrix} 1-0{,}6L & -0{,}8L \\ -0{,}2L & 1-0{,}6L \end{pmatrix} \begin{pmatrix} y_t \\ z_t \end{pmatrix} = \varepsilon_t$$

donde

$$A(1) = \begin{pmatrix} 0{,}4 & -0{,}8 \\ -0{,}2 & 0{,}4 \end{pmatrix} \Rightarrow \text{rango } A(1) = 1$$

Nótese a partir de (21) que

$$A(1) = B\alpha' = \begin{pmatrix} 0{,}4 \\ -0{,}2 \end{pmatrix} (1,-2)$$

Por tanto la representación ECM viene dada por

$$(1-L) \begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0{,}4 \\ -0{,}2 \end{pmatrix} (1,-2) \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \varepsilon_t$$

En este caso existe causalidad en ambas direcciones. Finalmente la representación TC viene dada por,

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0{,}5 & 1 \\ 0{,}25 & 0{,}5 \end{pmatrix} \begin{pmatrix} y_t^{(1)} \\ z_t^{(1)} \end{pmatrix} - \sum_{2}^{t} \left[\sum_{t-i+1}^{t-1} C_j \right] \varepsilon_i$$

$$\text{con } C_j = (0{,}2)^j \begin{pmatrix} -0{,}5 & 1 \\ 0{,}25 & -0{,}5 \end{pmatrix}$$

Nótese que en la representación TC, el producto interno $\alpha' x_t$ elimina las tendencias comunes $y_t^{(1)}$ y $z_t^{(1)}$

3.3. Extensiones de los Teoremas de Representación

En esta sección se extienden los teoremas de la sección anterior, con especial atención a la representación ECM, a casos en que el vector x_t tiene más de una raíz unitaria o posee raíces unitarias en las frecuencias estacionales. Con el fin de ilustrar la forma de proceder en casos más generales, consideremos la representación ECM en los casos de una doble raíz unitaria $(1-L)^2$ y de una raíz unitaria en la frecuencia semestral, $(1-L^2)$.

Ejemplo 5

Supongamos que $\{x_t\}$ tiene la representación de Wold

$$(1-L)^2 x_t = C(L)\,\varepsilon_t \tag{28}$$

donde C(L) puede desarrollarse del siguiente modo

$$C(L) = \gamma_1 + \gamma_2(1-L) + C^*(1-L)^2 \tag{29}$$

con υ_1=C(1), $\gamma_2 = -C^{(1)}(1)$ y $C^*(L)$ es invertible. Supongamos que existe un vector de cointegración α tal que

$$\alpha'\gamma_1 = \alpha'\gamma_2 = 0 \tag{30}$$

A partir de la representación VAR, se obtiene

$$A(L)\,C(L) = (1-L)^2\, I \Rightarrow A(1)\,C(1) = 0 \tag{31}$$

por lo que A(1) puede escribirse como

$$A(1) = B\alpha' \tag{32}$$

Derivando (31) con respecto a L y evaluando en L=1, se obtiene

$$A'(1)C(1) + A(1)\,C'(1) = 0 \tag{33}$$

Utilizando (30) y (32) se sigue que $A'(1)C(1) = 0$, de forma que $A'(1)$ puede escribirse en la forma

$$A'(1) = D\,\alpha' \tag{34}$$

Utilizando un argumento similar al usado en la obtención de (25), una expansión de A(L) similar a (29) conduce a la siguiente representación ECM

$$A^{**}(L)\Delta^2 x_t = -B\alpha' x_{t-1} - (B-D)\,\alpha'\Delta x_{t-1} + \varepsilon_t \tag{35}$$

con $A^{**}(0) = I$

Ejemplo 6

Supongamos que $\{x_t\}$ tiene la representación de Wold

$$(1-L^2)\, x_t = C(L)\,\varepsilon_t \tag{36}$$

La matriz C(L) puede escribirse como

$$C(L) = \gamma_1(1-L) + \gamma_2(1+L) + C^*(L)\,(1-L^2) \tag{37}$$

donde $\gamma_1 = C(-1)/2$, $\gamma_2 = C(1)/2$ y $C^*(L)$ es invertible

A partir de la representación VAR

$$A(L)\ C(L) = (1-L)\ (1+L)\ I \Rightarrow A(1)\ C(1) = A(-1)\ C(-1) = 0 \qquad (38)$$

Así pues, habrá cointegración en la raíz $(1+L)$ si existe un vector de cointegración α_1 tal que $\alpha'_1\ \gamma_1 = \alpha'_1\ C(-1) = 0$. De forma similar habrá cointegración en la raíz $(1-L)$ si existe un vector α_2 tal que $\alpha'_2\ \gamma_2 = \alpha'_2\ C(1) = 0$. Por tanto podemos expresar $A(1)/2 = B\alpha'_2$ y $A(-1)/2 = D\alpha'_1$. Utilizando el desarrollo (37) para $A(L)$ se obtiene la representación ECM

$$A^{**}(L)\ \Delta_2 x_t = B\alpha'_2\ \Delta x_{t-1} - D\alpha'_1\ (x_{t-1} + x_{t-2}) + \varepsilon_t \qquad (39)$$

con $A^{**}(0) = I$

Nótese que si no existe cointegración en $(1+L)$, $\alpha_1=0$ y el segundo término en el lado derecho de (39) desaparece. Igualmente si no hay cointegración en $(1-L)$, $\alpha_2=0$ y el primer término no aparece en (39).

Argumentos similares pueden utilizarse para obtener las representaciones ECM para otras frecuencias estacionales (en el caso de periodicidad trimestral, véase Engle et al. (1988) y Banerjee (1989)) e incluso cuando existen raíces en frecuencias regulares y estacionales (véase Osborn et al. (1989)).

4. ESTIMACION, CONTRASTACIÓN Y PREDICCION EN SISTEMAS COINTEGRADOS

En esta sección se discuten aspectos relativos a la estimación, contrastación y predicción en sistemas cointegrados. Con el fin de simplificar el análisis limitaremos la discusión a sistemas C(1,1), ofreciendo generalizaciones cuando sea necesario

4.1. Estimación de Vectores de Cointegración

Considérese el problema de estimar el vector de cointegración α en el siguiente modelo estático

$$\alpha' x_t = e_t \qquad (40)$$

Si todos los elementos de x_t son I(1), en general cualquier combinación de dichas variables será I(1), excepto en el caso de la combinación lineal dada por el vector α. Por tanto, casi todas las combinaciones lineales darán lugar a procesos e_t con varianza asintóticamente infinita. Dado que un procedimiento sencillo de estimación, como MCO, minimiza la varianza de e_t, el vector estimado $\hat{\alpha}$, después de normalizar a la unidad el coeficiente de una de las variables, representará una excelente aproximación al verdadero vector de cointegración α, si es que éste existe.

La sencillez del resultado anterior explica, en gran proporción, por qué el interés en el concepto de cointegración ha explotado como si se tratase de una serie integrada. Ello implica que para parametrizar una relación de equilibrio a largo plazo entre un conjunto de variables, todo lo que se necesita es estimar, mediante MCO, una regresión estática entre dichas variables. Esta regresión puede incluso estimarse en la primera etapa del proceso de especificación dinámica, tal como sugieren Engle y Granger (1987) en el procedimiento en "dos etapas" que discutirá posteriormente. En cualquier caso, dicha regresión puede utilizarse como un contraste inicial que indique la medida en que las predicciones de equilibrio sugeridas por la teoría económica están en consonancia con los datos disponibles. De esta manera se tendrá evidencia de si es fructífero emplear recursos posteriores en la modelización de la dinámica a corto plazo, alrededor de la relación de equilibrio, tal como ilustra la representación ECM.

De acuerdo con la discusión anterior, el estimador minimocuadrático de cualquier vector de cointegración convergerá rápidamente a su verdadero valor. Para analizar esta propiedad considérese el siguiente PGD

Ejemplo 7

$$y_t = \alpha z_t + e_t \; ; \; e_t = (1-\rho L)\, \xi_t \, ; |\rho| < 1$$

$$\Delta z_t = \varepsilon_t$$

$$\varepsilon_t \sim niid(0,\sigma_\varepsilon^2),\; \xi_t \sim niid(0,\sigma_\xi^2);\; E(\varepsilon_t \xi_s) = 0$$

El modelo que se estima es

$$y_t = \mu + \alpha x_t + u_t$$

lo que da lugar al siguiente conjunto de resultados

$$T(\hat{\alpha}-\alpha) \Rightarrow f_1(W) \; ; \; T(1-R^2) \Rightarrow f_3(W) \; ; \; t_{\alpha=0} \Rightarrow f_4(W)$$

$$T^{1/2}(\hat{\mu}) \Rightarrow f_2(W) \; ; \; DW \Rightarrow 2(1-\rho)$$

donde $f_i(W)$ representan genéricamente funcionales de procesos de Wiener del tipo (i)-(iii) en la Sección 2 (véase Stock (1987)). La interpretación de los resultados ilustra claramente la discusión informal previa. La pendiente estimador en la regresión estática converge a su verdadero valor α a velocidad $Op(T^{-1})$ en vez de $Op(T^{-1/2})$ como ocurre con series I(0). La intuición es clara ya que $\ddot{\alpha}$ se computa con el ratio de una covarianza, que de acuerdo con (iii) es $Op(T^{-1})$, respecto a una varianza, que de acuerdo con (i) es $Op(T^{-2})$, dado que tanto y_t como z_t son I(1). Por tanto el sesgo es $Op(T^{-1})$. Sin embargo, el t-ratio de $\hat{\alpha}$ tiene una distribución límite no degenerada que, en general, no coincide con la normal estandarizada por lo que los valores reportados de dicho estadístico carecen usualmente de sentido. Puesto que, en este caso, tanto y_t como z_t carecen de deriva (los componentes determinísticos han sido sustraídos), $\hat{\mu}$ converge consistentemente a cero, pero a una velocidad $Op(T^{-1/2})$. El coeficiente de correlación múltiple R^2 también es $Op(T^{-1})$ consistente de la unidad, reflejando el hecho de que en un modelo bivariante cointegrado existe un único vector de cointegración, por lo que el producto de las pendientes en las regresiones de y sobre z y de z sobre y, es la unidad. Esta característica será explotada posteriormente. Finalmente, el estadístico Durbin-Watson (DW) converge al resultado estándar, bajo el supuesto simplificador de una perturbación AR(1).

Un importante tema asociado a los resultados anteriores se refiere a la existencia de sesgos de simultaneidad o errores en las variables. Los sesgos que aparecen en dichas ocasiones, se derivan de las correlaciones existentes entre los regresores y las perturbaciones, que normalmente se suponen de orden Op(T). Sin embargo, dado que, por ejemplo, en el modelo del Ejemplo 7, $\Sigma\, z_t\, e_t$ tiene un orden probabilístico inferior al de $\Sigma\, z_t^2$, tales sesgos desaparecen asintóticamente. Ello implica que, términos de consistencia de los estimadores, ninguno de los problemas anteriores importa.

Ejemplo 8

Considérese el modelo canónico de errores en las variables

$$y_t = \beta\, z_t^* + u_t$$

$$z_t = z_t^* + \varepsilon_t$$

donde $E(z^*_t u_t) = E(z^*_t \varepsilon_t) = E(u_t \varepsilon_t) = 0$; $plim\ T^{-1}\ \Sigma x_t^2 = \sigma_x^2$; $plim\ T^{-1}\ \Sigma x_t^{*2} = \sigma_{x^*}^2$..

El estimador MCO de β es tal que

$$plim\ \hat{\beta} = \beta\ \sigma_{x^*}^2 / \sigma_x^2 < \beta$$

Por tanto $\hat{\beta}$ está sesgado a la baja y además no está identificado, puesto que $\sigma_{x^*}^2$ no puede estimarse a partir de los momentos muestrales. Maravall (1979) discute condiciones generales bajo las cuales la identificación de β es posible. Estas condiciones explotan ciertos rasgos de la dinámica del modelo. Por ejemplo si se conoce que

$$x^*_t = \rho\, x^*_{t-1} + v_t;\ |\rho| < 1;\ v_t \sim niid(0, \sigma_v^2)$$

$$E(v_t\, u_t) = E(v_t\, \varepsilon_t) = 0$$

entonces

$$plim\ \hat{\beta} = (1+\gamma)^{-1}\ \beta;\ \ \gamma = \sigma_\varepsilon^2(1-\rho^2)/\sigma_v^2$$

Puesto que x_t sigue el siguiente proceso ARIMA

$$x_t = \rho\, x_{t-1} + v_t + \varepsilon_t - \rho\, \varepsilon_{t-1}$$

entonces ρ, σ_ε^2 y σ_v^2 pueden estimarse consistentemente a partir de los momentos muestrales γ_{xx}, γ_{xx-1} y γ_{xx-2}. Sin embargo, si x^*_t es I(1), i.e. ρ=1, entonces γ=0 y β puede ser identificado únicamente a partir de $\hat{\beta}$, que es consistente.

Este resultado tiene implicaciones importantes para determinados argumentos que abundan en textos de macroeconomía. Así, por ejemplo, Friedman (1957) explica la divergencia existente entre las estimaciones de las propensiones media y marginal a consumir en base a un argumento basado en "errores en las variables", i.e. la confusión entre renta actual y permanente. Si la renta sigue un proceso integrado este argumento no es válido, tal como se ha visto previamente. Stock (1988), aplicando la teoría asintótica de procesos de Wiener, demuestra que la propensión marginal a consumir presenta un sesgo negativo de 0,15 en la regresión estática de consumo sobre renta. Ajustando dicho sesgo, demuestra que ambas propensiones son aproximadamente igual a la unidad.

El resultado más importante de esta sección se refiere a la "super consistencia" de $\hat{\alpha}$. Sin embargo, tal como se apuntaba en el párrafo anterior, los sesgos, a pesar de ser $Op(T^{-1})$, pueden ser importantes en muestras finitas. Supóngase que los ratios de convergencia de dos estimadores fuesen $T^{-1/2}$ y $10^{10}\ T^{-1}$, en este caso se necesitarían muestras muy grandes para que el segundo estimador dominase al primero. En un amplio estudio de Monte-Carlo, Banerjee et al. (1986) examinan los sesgos derivados de la regresión estática. Dado que el coeficiente de correlación múltiple R^2 converge a la unidad también a velocidad $Op(T^{-1})$, proponen $(1-R^2)$ como una aproximación al sesgo. Por ejemplo si se supone un PGD como en el Ejemplo 7, entonces la relación entre sesgo y R^2 resulta ser

$$\hat{\alpha} - \alpha = \alpha(1-R^2) + Op(T^{-1}) \tag{41}$$

Si T es suficientemente grande, α=1 y R^2=0.90, el valor de $\hat{\alpha}$ estará alrededor de 0.9, lo que implica un sesgo apreciable. Este resultado sugiere que las relaciones de cointegración estimadas en modelos estáticos con R^2 no muy cercanos a la unidad, deben interpretarse con escepticismo (véase, por ejemplo, Campbell y Shiller (1987), donde el R^2 de la relación entre tipos de interés a corto y largo plazo es 0.8). Sin embargo, resulta importante destacar que en el contexto de regresión múltiple, el R^2 de una ecuación siempre aumenta cuando se introduce una variable adicional, por lo que un R^2 alto no es condición necesaria para la obtención de sesgos reducidos.

En el Ejemplo 7, el PGD es esencialmente estático, con lo que el modelo reproduce las características del PGD. En general, cuando la relación entre y_t y z_t sea dinámica, el modelo estático traslada toda la dinámica al residuo u_t. Los términos dinámicos pueden reparametrizarse en términos de variable I(0), en la forma Δy_{t-i}, Δz_{t-i} y $(y-\alpha z)_{t-i}$ donde los valores de i dependerán de la naturaleza del proceso ARIMA multivariante que genere $x_t=(y_t, z_t)'$, tal como se muestra en la representación ECM de un modelo C(1,1) dada en (25). Con el fin de ofrecer una simple ilustración de este hecho sustitúyase la primera ecuación en el PGD del Ejemplo 7 por

$$y_t = \alpha_1 y_{t-1} + \alpha_2 z_t + \alpha_3 z_{t-1} + \varepsilon_t;\ \varepsilon_t \sim niid(0,\sigma_\varepsilon^2) \tag{42}$$

La expresión anterior puede reparametrizarse como

$$y_t = \alpha z_t + e_t \tag{43}$$

donde $\alpha = (\alpha_2+\alpha_3)(1-\alpha_1)^{-1}$ y $e_t = -(1-\alpha_1)^{-1}[\alpha_1 \Delta y_t + \alpha_3 \Delta z_t - \varepsilon_t]$

Puesto que Δy_t y Δz_t son I(0), e_t es I(0) y contiene correlación serial. Phillips (1988) demuestra que es precisamente la presencia de correlación serial en e_t, el motivo de obtener sesgos en la estimación de α en muestras finitas. Una alternativa obvia es la estimación de α directamente a partir de (42) (cuando z_t sea débilmente exógena, como en el presente caso), obtenido a partir del ratio $\hat{\alpha}=(1-\hat{\alpha}_1)^{-1}(\hat{\alpha}_1+\hat{\alpha}_3)$. En este caso la perturbación ε_t carece de correlación serial. Banerjee et al. (1986) comparan, mediante simulación, los sesgos que se obtienen por ambos métodos, regresión estática y dinámica, obteniendo una clara ventaja del método dinámico.

Finalmente, otra importante implicación que se deriva de la existencia de cointegración, es que, en contraste con el análisis estándar de regresión donde la multicolinealidad se considera un problema, en el contexto de sistemas cointegrados dinámicos, su existencia es esencial. Por ejemplo, considérese el modelo (42) reparametrizado como

$$\Delta y_t = \alpha_2 \Delta z_t + (\alpha_1 - 1) y_{t-1} + (\alpha_2 - \alpha_3) z_{t-1} \tag{44}$$

Típicamente y_{t-1} y z_{t-1} serán altamente colineales y consecuentemente sus t-ratios serán bajos. Eliminar cualquiera de las dos variables sería una solución desastrosa al problema de multicolinealidad, puesto que en ese caso el regresando sería I(0) y el conjunto de los regresores, I(1). Siendo I(1), los segundos momentos de y_{t-1} y z_{t-1} son muy grandes relativamente a los de Δz_t y, por tanto, cuando la matriz (X'X) se invierta, la submatriz que incluye y_{t-1} y z_{t-1} tendrá una "quasi-singularidad" correspondiente al vector de cointegración $(1,-\alpha)$. Esto, sin embargo, no es un problema, pues todo lo que se necesita es parametrizar (44) en la forma ECM, esto es

$$\Delta y_t = \alpha_2 \Delta z_t + (\alpha_1 - 1)(y_{t-1} - \alpha z_{t-1}) + \varepsilon_t \tag{45}$$

donde los regresores son ahora mucho menos colineales. En este sentido, la presencia de multicolinealidad es un signo de la existencia de cointegración en la relación estimada.

4.2. Procedimientos de estimación en dos etapas: Estática y Dinámica

Una vez que el vector de cointegración α ha sido estimado a partir de una regresión estática del tipo (43), Engle y Granger (1987) sugieren sustituir $\hat{e}_{t-1}$ en la representación ECM, por ejemplo (45), y estimar en una segunda etapa los parámetros α_1 y α_2 asociados a las variables I(0). La intuición subyacente a este procedimiento es que el estimador $\tilde{\alpha}$, en la primera etapa, es Op(T^{-1}), mientras que los estimadores $\tilde{\alpha}_1$ y $\tilde{\alpha}_2$, en la segunda etapa, son Op($T^{-1/2}$). Engle y Granger demuestran que los estimadores en dos etapas de los coeficientes de una ecuación perteneciente a un sistema ECM, utilizando $\hat{\alpha}$ como verdadero valor, tienen idénticas distribuciones límite a las obtenidas a partir de estimadores máximoverosímiles que utilicen el verdadero valor de α. Aplicaciones recientes de esta metodología de estimación, se encuentra, inter alia, en los trabajos de Hall (1986), Jenkinson (1986), Campbell y Shiller (1987), Campbell (1987) y Andrés et al. (1989).

Tal como se mencionaba previamente, un procedimiento alternativo de estimación en dos etapas, consiste en estimar directamente $\hat{\alpha}$ a partir de la relación dinámica, del tipo (42), obteniendo $\tilde{\alpha}$. Banerjee et al. (1988) demuestran los estimadores obtenidos por este procedimiento tienen las mismas propiedades que los anteriores. Además de la posible reducción en sesgo de este segundo procedimiento (véase Banerjee et al. (1986)), cuando z_t es débilmente exógena, es posible demostrar (véase Dolado y et al. (1989b)) que $T(\hat{\alpha}-\alpha)$ es asintóticamente normal mientras que $T(\tilde{\alpha}-\alpha)$ no es normal, cuando el modelo tiene dinámica. Puesto que muchos modelos de regresión pueden interpretarse como distribuciones condicionales autónomas o, en otras palabras, los regresores en dichos modelos son débilmente exógenos, en estos casos puede llevarse a cabo inferencia sobre los coeficientes de las variables que aparecen en la relación de cointegración.

Como una crítica adicional del primer procedimiento en dos etapas, Gourieroux et al. (1987) examinan las velocidades de convergencia de $\hat{\alpha}$ y $\hat{\hat{\alpha}}$, bajo exogeneidad débil, cuando las variables son I(d). Se analizan tres casos de no estacionariedad, incluyendo diferenciación fraccional, i.e. $0{,}5 < d < 1$, d=1 y $d > 1$, encontrando que los grados de convergencia de $\hat{\alpha}$ son $T^{-(2d-1)}$, T^{-1} y T^{-1}, mientras que los correspondientes a $\hat{\hat{\alpha}}$ son T^{-d} en todos los casos [6]. A la vista de estos resultados, resulta claro que $\hat{\hat{\alpha}}$ converge más rápidamente que $\hat{\alpha}$ a su verdadero valor en todos los casos, excepto cuando d=1. Dado que d=1 es un caso tan especial, de ello se deriva cierta desconfianza hacia el procedimiento estático en dos-etapas.

Finalmente es necesario destacar la importancia de imponer las raíces unitarias, que se conoce que existen a fin de llevar a cabo inferencia sobre el vector de cointegración bajo la hipótesis de cointegración. Tal como se mencionaba anteriormente, Sims et al. (1986) han demostrado la posibilidad de aplicar técnicas estándar de contrastación en modelos VAR con regresores integrados y en presencia de cointegración. Sin embargo, este resultado no se verifica cuando se trata de hacer inferencia sobre los parámetros del vector de cointegración (véase Phillips (1988) y Dolado et al. (1989b)), tal como se ilustra en el siguiente ejemplo

Ejemplo 9

Considérese el siguiente PGD

$$y_t = \beta x_t + \varepsilon_{1t}$$

$$x_t = \alpha x_{t-1} + \varepsilon_{2t};\ \alpha = 1;\ x_0 = 0,\ \varepsilon_i \sim niid(0,\sigma_i^2);\ E(\varepsilon_{1t}\,\varepsilon_{2s}) = \rho\,\sigma_1\,\sigma_2\,\delta_{ts}$$

Supongamos inicialmente que se impone α=1. Como las perturbaciones están correlacionadas, β ha de estimarse por máxima verosimilitud, que corresponde a estimar β por MCO en el modelo ampliado

$$y_t = \beta\, x_t + \gamma\, \Delta\, x_t + w_t$$

donde $\gamma = \rho\sigma_1/\sigma_2$ y $E[w_t \varepsilon_{2t}]=0$. Puede, entonces, demostrarse que, de acuerdo con (i)-(iii) en la Sección 2

$$T(\hat{\beta}-\beta) \Rightarrow \frac{\sigma_w \int_0^1 V(r)\, dW(r)}{\sigma_v \int_0^1 V^2(r)\, dr} \equiv MN$$

donde $P_t(=\sum_1^t \varepsilon_{2i}) \Rightarrow V(r)$ y $S_t(=\sum_1^t w_i) \Rightarrow W(r)$. Dada que, por construcción, V y W son procesos de Wiener independientes, es posible demostrar que el funcional de Wiener anterior sigue una mixtura de distribuciones normales (MN) (véase Phillips and Park (1988)) que es equivalente a la distribución $N(0,\sigma_2^2 \sigma_w^2 / 2)$.

Sin embargo, si $\alpha=1$ no se impone, sino que α se estima, entonces es posible demostrar que el estimador maximoverosímil de β en este caso, es tal que

$$T(\hat{\beta}-\beta) \Rightarrow \gamma \frac{1/2[V(1)^2-1]}{\int_0^1 V(r)^2 dr} + \frac{\sigma_w \int_0^1 V(r)\, dW(r)}{\sigma_v \int_0^1 V^2(r)\, dr}$$

Puesto que el primer término de la distribución asintótica corresponde a (4) (con $\sigma_\varepsilon^2=\sigma^2$), se conoce que esta distribución no es estándar. Por tanto, a menos que $\gamma=0$ ($\rho=0$) la inferencia sobre β ha de hacerse con distribuciones simuladas.

4.3. Contrastación

Contrastar la existencia de cointegración entre un conjunto de variables corresponde a contrastar la existencia de una raíz unitaria. Por tanto, el análisis subyacente a este tipo de contraste, tal como se apunta brevemente en la Sección 2, se aplica también a este caso. La diferencia estriba en que ahora se contrasta la presencia de una raíz unitaria en una serie construida, i.e. residuos minimocuadráticos, en vez de series originales. Por este motivo, los valores críticos de este tipo de contrastes han de ser ajustados al alza, ya que de otra forma el error de tipo 1 estaría exagerado. Si no se rechaza la existencia de raíces unitarias en dichos residuos, las "desviaciones del equilibrio" serían integradas y por tanto no existiría un mecanismo de ajuste sistemático del sistema hacia la solución de equilibrio.

En las aplicaciones empíricas los contrastes de mayor uso son el contraste de cointegración basado en el estadístico Durbin-Watson (CRDW) y los contrastes de Dickey y Fuller (1979) en sus versiones simples (DF) y aumentado (ADF). El contraste CRDW, sugerido por Sargan y Bhargava (1983) está basado en la definición usual del estadístico DW

$$CRDW = (\Sigma\, \hat{e}_t^2)^{-1}\, [\Sigma(\hat{e}_t - \hat{e}_{t-1})^2] \qquad (46)$$

donde $\hat{e}_t$ son los residuos minimocuadráticos de (40). Existen diversos problemas con este contraste. En efecto, mientras que cuando se trata de contrastar el grado de integración de una serie (con una constante como único regresor), los valores críticos del contraste, computados por Sargan y Bhargava, son exactos, el mismo contraste aplicado al caso de cointegración depende del número de regresores en la relación de cointegración, existiendo sólo cotas de la región crítica. Esto ocurre porque, como en el caso estandar de aplicación del DW (basado en la hipótesis nula de perturbaciones incorrelacionadas), los valores críticos exactos del CRDW (basado en la hipótesis de una raíz unitaria en las perturbaciones) son una función del PGD. Las cotas de este contraste proporcionan un intervalo que puede utilizarse para no rechazar la hipótesis nula de ausencia de cointegración, pero la amplitud de dicho intervalo aumenta conforme lo hace el número de regresores. Sin el uso, por ejemplo, de la Rutina de Imhof (1961), es difícil hacer inferencia cuando el valor del CRDW cae dentro del intervalo. Por supuesto, resulta factible computar valores críticos exactos para un PGD dado. Así, Engle y Granger (1987) presentan valores críticos para un paseo aleatorio bivariante, pero dichos valores no son generalmente aplicables a otros experimentos cuyos PGD sean más complejos. En este sentido el CRDW resulta ser *asintóticamente similar* (véase Cox y Hinkley (1974)) cuando el verdadero proceso seguido por las perturbaciones en un AR(1).

Problemas similares afectan al estadístico DF, pero no afectan al ADF si el número de desfases Δe_{t-i} que aparecen en el contraste coincide con el del PGD. En la práctica el número de desfases que se utiliza es, en gran medida, ad-hoc y ello afecta a la potencia del contraste. Todo ello sugiere que quizá se deba prestar mayor atención a las versiones no-paramétricas (véase Ouliaris et al. (1988)) o recursivas y secuenciales (véase Dolado et al. (1989a)) de dichos contrastes.

Con fines ilustrativos se presenta en el Cuadro 1, los valores críticos (5%) de los tres contrastes anteriores, computados por Engle y Yoo (1987), donde el PGD es un paseo aleatorio n-variante, con $n \leq 5$

CUADRO 1

Valores Críticos del Contraste de Cointegración (5%)

n	T	CRDW	DF	ADF
	50	0,78	-3,67	-3,29
2	100	0,39	-3,37	-3,17
	200	0,20	-3,37	-3,25
	50	0,89	-4,11	-3,75
3	100	0,45	-3,93	-3,62
	200	0,30	-3,78	-3,78
	50	1,10	-4,35	-3,98
4	100	0,55	-4,22	-4,02
	200	0,40	-4,18	-4,13
	50	1,10	-4,76	-4,15
5	100	0,66	-4,58	-4,36
	200	0,47	-4,48	-4,43

Nota: Los valores críticos del contraste CRDW han sido tomados de Engle y Granger (1987) para n=2. Los restantes valores han sido interpolados utilizando los resultados en Sargan y Bhargava (1983). Los valores críticos de los contrastes DF y ADF han sido tomados de Engle y Yoo (1987).

Tal como ocurre con los contrastes de raíces unitarias en series temporales individuales es importante destacar la falta de potencia de estos contrastes cuando la verdadera raíz se encuentra en un entorno de la unidad. Así, por ejemplo, Engle y Granger muestra que cuando el PGD de las perturbaciones de la relación de cointegración es un AR(1) con parámetro igual a 0,9, las potencias de los contrastes CRDW, DF y ADF (al 5%) son 20%, 15% y 11% respectivamente. Cuando se altera el PGD a un proceso AR más general la potencia del ADF es del 60%, dominando a los contrastes CRDW y DF. A la vista de estos resultados quizá resulte conveniente aconsejar el uso del CRDW en una primera instancia, dada su simplicidad, y, en caso de incertidumbre sobre la posibilidad de rechazo de la hipótesis nula, utilizar el ADF [7].

Un enfoque alternativo, discutido previamente, se basa en contrastar la presencia de cointegración utilizando la solución de equilibrio a largo plazo del modelo dinámico, como en (42), especialmente cuando los regresores sean débil o fuertemente exógenos. En la representación ECM, como en (45), el t-ratio del término de corrección del error, $(\hat{\alpha}_1-1)$, es un estadístico de interés, ya que si no existiese cointegración debería converger a cero a velocidad $Op(T^{-1})$. Banerjee et al. (1986) muestran, por simulación como dicho t-ratio tiene el tamaño adecuado al 5%, aunque los resultados de Dolado et al. (1989b) muestran que puede haber divergencias para otros PGD's. Sin embargo, de acuerdo con el estudio de Banerjee et al., la potencia de este contraste es muy superior a la de los contrastes basados en la regresión estática. Una posible explicación de esta ventaja comparativa reside en los sesgos en muestras finitas que presenta el enfoque estático.

Finalmente, resulta importante comentar que, dada la fragilidad de los contrastes, otros procedimientos auxiliares menos refinados estadísticamente puede resultar útiles. La observación del correlograma de los residuos puede ayudar en gran medida. Granger y Weiss (1983) también sugieren aumentar los coeficientes estimados del vector de cointegración en una cierta proporción, digamos el 10%, y examinar, entonces, si la suma de cuadrados de las desviaciones del nuevo vector es bastante superior a la obtenida con el vector estimado. La intuición de este procedimiento informal es clara, puesto que sólo el uso del verdadero vector de cointegración hará que dicha varianza sea finita y, por tanto, será fácilmente distinguible de otros casos.

4.4. Predicción con Sistemas Cointegrados

Tal como han demostrado Engle y Yoo (1987), las predicciones derivadas de sistemas cointegrados tienen una propiedad que no comparte los sistemas de variables integradas. Esta propiedad implica la existencia de combinaciones lineales de dichas predicciones que son idénticamente cero para horizontes temporales largos, independientemente del punto de origen y, además, la varianza del error de predicción de dichas combinaciones es finita, mientras que, para cualquier otra combinación, diverge conforme se amplie el horizonte de predicción. El siguiente ejemplo ilustra esta propiedad.

Ejemplo 10

Considérese el proceso $\{x_t\}$ cuyo PGD es

$$\Delta x_t = \varepsilon_t - C_1 \varepsilon_{t-1};\ \varepsilon_{t-1};\ \varepsilon_0 = x_0 = 0;\ \varepsilon_t \sim (0,\Omega)$$

tal que x_t es C(1,1). Por tanto existe un vector α, tal que

$$\alpha' C(1) = \alpha'(I - C_1) = 0$$

El valor x_{t+h} puede expresarse recursivamente como

$$x_{t+h} = \varepsilon_{t+h} + (\text{I-C}_1) \sum_1^{t+h-1} \varepsilon_j$$

Por tanto

$$\text{E}(x_{t+h}/t) = x_{t+h/t} = (\text{I-C}_1) \sum_1^{t} \varepsilon_j$$

y

$$e_{t+h/t} = \varepsilon_{t+h} + (\text{I-C}_1) \sum_1^{h} \varepsilon_{t+j}$$

Se sigue que

$$\alpha' \, x_{t+h/t} = 0$$

y

$$\text{var}\,(\alpha' \, e_{t+h/t}) = \alpha' \, \Omega \, \alpha < \infty$$

Con el fin de examinar el comportamiento predictivo de métodos de estimación que ignoran la cointegración existente y aquellos que la tienen en cuenta. Engle y Yoo analizan el comportamiento predictivo relativo a un sistema cointegrado estimado como un VAR irrestringido (equivalente a aplicar MCO ecuación a ecuación) y reparametrizado como un modelo ECM, aplicando el procedimiento estático en dos etapas El PGD viene dado por

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0{,}8 & 0{,}1 \\ 0{,}8 & 0{,}6 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \varepsilon_t \tag{47}$$

donde ambas variables son I(1) y el vector de cointegración es (1,-0,5).

Los resultados, medidos a través del Error Cuadrático Medio de la Predicción (ECMP) muestran una ligera ventaja del método VAR hasta un horizonte de 4 períodos, probablemente debido al sesgo en la estimación del vector cointegrador, pero posteriormente el método ECM se muestra muy superior. Así, después de 20 períodos, el ECMP del modelo ECM es un 40% inferior al del VAR, y probablemente esa mejora en la predicción sería muy superior si el vector de cointegración se hubiese estimado con menor sesgo.

5. ESTIMACION Y CONTRASTACION EN UN CONTEXTO DE MAXIMA VEROSIMILITUD

El propósito de esta sección es el de ofrecer un resumen de ciertos desarrollos recientes (véase Johansen (1988) y Johansen y Juselius (1988)) que tratan los problemas de estimación y contrastación de sistemas cointegrados en un contexto de máxima verosimilitud [8]. El enfoque está limitado, por el momento, a sistemas C(1,1) que admiten la representación VAR dada por

$$x_t = \sum_{1}^{p} A_i x_{t-i} + \varepsilon_t \; ; \; \varepsilon_t \sim NI(0,\Omega) \qquad (48)$$

La hipótesis nula de interés es que la dimensión del subespacio de vectores de cointegración es r, una hipótesis que, en este contexto, viene expresada mediante una simple formulación paramétrica en términos de los vectores de cointegración y sus ponderaciones respectivas. Se utiliza un procedimiento de máxima verosimilitud para contrastar dicha hipótesis.

El procedimiento se basa en la reparametrización de (48) en términos de

$$\Delta x_t = -A(1)\, x_{t-1} + \sum_{1}^{p-1} \Pi_i \Delta x_{t-p+1} + \varepsilon_t \qquad (49)$$

donde $A(1) = I - \sum_{1}^{p} A_i$ y $\Pi_i = - \sum_{i+1}^{p} A_i$

Si el sistema (49) está cointegrado, se conoce, a partir de (21), que $A(1)=B\Gamma'$. La hipótesis es, por tanto, H_0: rango $A(1)=r$. Existen tres casos de interés. En primer lugar, si $r=n$ entonces la matriz A(1) tiene rango completo por lo que x_t es estacionario. En segundo lugar, si $r=0$ entonces no existe cointegración y el modelo debería ser estimado en diferencias, eliminando el término en niveles en (49). En tercer lugar si $0<r<n$, existen r vectores de cointegración.

La función de verosimilitud (en logs) de (49) puede escribirse del siguiente modo

$$L(B,\Gamma,\Pi_i;\Omega) = -\frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_{1}^{T}[\Delta x_t + B\Gamma' x_{t-1} - \Sigma\, \Pi_i \Delta x_{t-i+1}]' \Omega^{-1}$$
$$[\Delta x_t + B\Gamma' x_{t-1} - \Sigma\, \Pi_i \Delta x_{t-i+1}] + \text{const} \qquad (50)$$

A continuación se concentra la función L respecto a las matrices Π_i's. Como es bien conocido del análisis de regresión multivariante, ello se logra mediante la obtención de las matrices de residuos de las regresiones de Δx_t y x_{t-1} sobre $\Delta x_{t-1} \ldots \Delta x_{t-p+1}$, denotando dichas matrices para R_{0t} y R_{1t} respectivamente. Seguidamente se fórmula el siguiente modelo de regresión concentrado.

$$R_{ot} = -B\Gamma' R_{1t} + \varepsilon_t \tag{51}$$

Dado que (51) es no lineal en las matrices B y Γ, sólo es posible concentrar respecto a B si Γ se toma como dada. Por tanto, los regresores en (51) habrán de ser interpretados como $\Gamma' R_{1t}$. Las soluciones obtenidas para B y Ω vienen dadas por

$$B(\Gamma) = -S_{01}\Gamma[\Gamma' S_{11} \Gamma]^{-1} \tag{52}$$

y

$$\Omega(\Gamma) = S_{00} - S_{01} \Gamma[\Gamma' S_{11}\Gamma]^{-1} \Gamma' S_{10} \tag{53}$$

tal que $S_{ij} = T^{-1} \sum_{1}^{T} R_{it} R'_{jt}$ $(i,j=0,1)$

Por tanto, la función de verosimilitud concentrada viene dada por

$$L^*(\Gamma) = \text{const} -T/2 \log |\Omega(\Gamma)| \tag{54}$$

Maximizando (54) respecto a r es equivalente a

$$\min_{\Gamma} | S_{00}-S_{01}\Gamma(\Gamma' S_{11}\Gamma)^{1} \Gamma' S_{10}| \tag{55}$$

A partir de la siguiente conocida relación entre determinantes de matrices particionadas (véase Dhrymes (1978))

$$\begin{vmatrix} S_{00} & S_{01}\Gamma \\ \Gamma' S_{10} & \Gamma' S_{11}\Gamma \end{vmatrix} = | S_{00}| \, | \Gamma' S_{11} \Gamma - \Gamma S_{10} S_{00}^{-1} S_{01} \Gamma | =$$

$$= | \Gamma' S_{11} \Gamma| \, | S_{00}-S_{01}\Gamma(\Gamma' S_{11}\Gamma)^{-1}\Gamma' S_{10}| \tag{56}$$

se sigue que minimizar (55) es equivalente a

$$\min_{\Gamma} | \Gamma' S_{11}\Gamma|^{-1} | \Gamma' S_{11}\Gamma-\Gamma' S_{10} S_{00}^{-1} S_{01}\Gamma| \tag{57}$$

Puesto que la solución al problema (57) carece claramente de unicidad, se impone la normalización $\Gamma' S_{11}\Gamma = I$, convirtiendo el problema en

$$\min_{\Gamma} | I-\Gamma' S_{10} S_{00}^{-1}\Gamma| \quad \text{s.a. } \Gamma' S_{11}\Gamma = I \tag{58}$$

La solución al anterior problema es bien conocida en la literatura (véase, por ejemplo, Anderson (1984)) y se obtiene calculando los valores propios de la matriz $S_{10} S_{00}^{-1} S_{01}$ con respecto a la matriz S_{11}, resolviendo

$$| \hat{\lambda} S_{11} - S_{10} S_{00}^{-1} S_{01} | = 0 \tag{59}$$

Los vectores propios correspondientes vienen dados por

$$S_{10} S_{00}^{-1} S_{01} \hat{v}_i = \hat{\lambda}_i S_{11} \hat{v}_i \; (i=1 \ldots n) \tag{60}$$

$$\text{s.a. } \hat{v}_i' S_{11} \hat{v}_j = S_{ij}$$

donde δ es la delta de Kroenecker.

Resulta posible demostrar que los valores propios $\hat{\lambda}_i \varepsilon(0,1)$ y que los vectores de cointegración agrupados en la matriz Γ, corresponden a los r vectores propios, $\hat{v}_i (i=1 \ldots r)$ asociados con los r valores propios mayores. Nótese que los estimadores maximoverosímiles de Π_i $(i=2 \ldots p)$ y de B y Ω pueden obtenerse a partir de (49), (52) y (53) sustituyendo Γ por su estimador $\hat{\Gamma}$.

Con el fin de construir un contraste basado en el ratio de verosimilitud (LR), se sustituye (56) en (54), obteniendo

$$L^*(\Gamma) = cnst - T/2 \; log \; \| \Gamma' S_{11} \Gamma |^{-1} (| S_{00} \| \Gamma' S_{11} \Gamma - \Gamma' S_{10} S_{00}^{-1} S_{01} \Gamma |) | \tag{61}$$

A continuación, utilizando (59) y la propiedad que establece que el determinante de una matriz es igual al producto de sus valores propios, bajo la hipótesis nula de que existen r vectores de cointegración, la función L* toma el valor

$$\tilde{L}^*(\Gamma) = -\frac{T}{2} \log | S_{00} | -\frac{T}{2} \sum_{1}^{r} \log (1-\hat{\lambda}_i) + \text{const} \tag{62}$$

donde los valores propios $\hat{\lambda}_i$ han sido ordenados de mayor a menor $(\hat{\lambda}_1 > \ldots > \hat{\lambda}_r)$. El máximo valor de la función, $\hat{L}^*$, correspondiente a la hipótesis no restringida, es igual a (62) con el sumatorio de 1 a n. por tanto el contraste LR de H_0: rango $\Gamma = r$ frente a H_1: rango $\Gamma = n$ viene dado por

$$LR = 2 \; (\hat{L}^* - \tilde{L}^*) = - T \sum_{r+1}^{n} \log (1-\hat{\lambda}_i) \tag{63}$$

Johansen (1987) demuestra que la distribución de LR no es estandar, correspondiendo a la versión multivariante del cuadrado de (i) en la Sección 2, dada por la siguiente traza (tr)

$$tr[\textstyle\int_0^1 W(r) dW(r)' (\int_0^1 W(r) W'(r) dr)^{-1} \int_0^1 dW(r) W'(r)] \tag{64}$$

CUADRO 2
Valores Críticos del Contraste LR (5%)

r	LR
0	57,2
1	38,6
2	23,8
3	12,0
4	4,2

NOTA: Los valores críticos han sido tomados de Johansen y Juselius (1988).

También es posible considerar hipótesis acerca de restricciones lineales en B y Γ bien separada o conjuntamente. Las hipótesis nulas acerca de B y Γ tiene la forma H_0: $B=H\gamma$ y $H_0'\Gamma=G\Psi$ mientras que la hipótesis conjunta será H''_0:$B\Gamma'=H\gamma\Psi'G'$. Las matrices H y G son $(n\times s)$ y $(n\times q)$ respectivamente, reduciendo el número de parámetros en B a $\gamma(s\times r)$ y en Γ a $\Psi(q\times r)$ donde $r<s<n$ y $r<q<n$. Todas estas hipótesis suponen restricciones sobre la matriz. A(1) que contiene n^2 parámetros $(nr+(n-r)r$, bajo la hipótesis de cointegración), mientras que bajo H_0, H'_0 y H'_0 contiene $sr+(n-r)r$, $qr+(n-r)$ y $qr+(s-r)r$ respectivamente Johansen y Juselius (1988) demuestran que la formulación de Wald de estos contrastes se distribuyen asintóticamente como distribuciones Ji-cuadrado con $(n-s)r$, $(n-q)r$ y $(2n-q-s)r$ grados de libertad respectivamente. la intuición subyacente a estos resultados es similar a la presentada en el Ejemplo 9. A este respecto, es posible demostrar que los resultados de Dolado et al. (1989b) acerca de la distribución asintótica normal de Γ, son fácilmente generalizables en este contexto, suponiendo que Ω es diagonal y que si $b_{ij}\neq 0$, entonces b_{ik} y b_{kj} $(k=1\ldots n)$ son cero. En otras palabras, se elimina la posibilidad de que aparezca una misma relación de cointegración en más de una ecuación, lo que unido a la diagonalidad de Ω, implica la existencia de regresores débilmente exógenos respecto a los elementos de Γ.

Finalmente, resulta interesante hacer notar la similitud del anterior procedimiento con el procedimiento de "tendencias comunes" desarrollado por Stock y Watson (1988b). Este último procedimiento permite determinar el número de vectores de cointegración por medio del contraste de la hipótesis nula de que el vector x_t presenta $(n-r)$ raíces unitarias distintas, frente a la alternativa de que tiene $(n-m)$ $(m>r)$ raíces unitarias. El contraste se construye transformando x_t tal que los primeros r componentes corresponden a los componentes estacionarios, mientras que los restantes (n-r) correspondan a componentes integrados. Bajo la hipótesis nula, los valores propios de la autorregresión de primer orden de los (n-r) componentes deberán ser igual a la unidad. Bajo la hipótesis alternativa, solamente (n-m) valores propios habían de ser igual a la unidad, de forma que el valor propio (n-m+1)-*smo* deberá ser inferior a la unidad. El estadístico del contraste viene dado por $T(\hat{\lambda}_{n-m+1}-1)$ y sus valores críticos han sido obtenidos por simulación, encontrándose en el artículo de Stock y Watson.

Notas

1. Como muestra del interés que ha despertado el tema podemos citar que, desde 1986, han aparecido en la literatura dos números especiales de revistas dedicadas al tema (véase Oxford Bulletin of Economics and Statistics (1986) y Journal of Economic Dynamics and Control (1988)) y tres panorámicas (véase Stock y Watson (1988a), Pagan y Wickens (1989) y Dolado, Jenkinson y Sosvilla-Rivero (1989).

2. El interés en el tema se remonta a nombres tan ilustres como Jevons (1884), Working (1934) y principalmente Yule (1926).

3. A lo largo de este artículo se restringe la discusión al concepto de *integración en varianza.* Un análisis excelente de dicho concepto aplicado a otros momentos se encuentra en Escribano (1987).

4. Una discusión extensa de las propiedades asintóticas de procesos $I(d)$ con $d > 1$ puede encontrarse en Gourieroux et al. (1987).

5. Una ampliación de los Teoremas de Representación a procesos no lineales puede encontrarse en Escribano (1986).

6. Para un análisis del concepto de diferenciación fraccional véase Granger y Joyeux (1980).

7. Bajo una interpretación similar puede llevarse a cabo contrastes de cointegración en diferentes frecuencias (véase Engle et al. (1988) y Osborn et al. (1989)).

8. Este enfoque está claramente relacionado con el enfoque de análisis canónico propuesto por Box y Tiao (1981) y Peña y Box (1987).

REFERENCIAS

ANDERSON, T. (1984). *"An Introduction to Multivariate Statistical Analysis"*, New York, Wiley.

ANDRES, J., ESCRIBANO, A. MOLINAS, C. y TAGUAS, D. (1989). "Los Determinantes de la Inversión Productiva Privada en España" Moneda y Crédito 1, 67-97.

AOKI, M. (1987). "An Alternative Measure of Random Walk Components in Time Series" *Economics Letters*, 24, 227-230.

BANERJEE, A., DOLADO, J., HENDRY, D. and SMITH, G. (1986). "Exploring Equilibrium Relationshisps in Econometrics through Static Models: Some Monte Carlo Evidence" *Oxford Bulletin of Economics and Statistics* 48, 253-277.

BANERJEE, A. and DOLADO, J. (1988). "Tests of the Life Cycle Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small-Sample Interpretations" *Oxford Economic Papers*, 40, 610-633.

BANERJEE, A., GALBRAITH, J. and DOLADO, J. (1988) "Dynamic Specification with the General Error-Correction Form" (de próxima aparición en *Oxford Bulletin of Economics and Statistics)*.

BANERJEE, A., DOLADO, J., HENDRY, D. and GALBRAITH, J. (1989) *"Equilibrium Error Correction and Cointegration in Econometrics"* (de próxima aparición en Oxford University Press).

BILLINGSLEY, P. (1986). *"Convergence of Probability Measures"* New York, Wiley.

BOX, G. and JENKINS, G. (1976). *"Time Series Analysis: Forecasting and Control"*; San Francisco, Holden-Day.

BOX, G. and TIAO, G. (1981). "A Canonical Analysis of Multiple Time Series with Applications" *Biometrika*, 64, 355-365.

CAMPBELL, J. (1987). "Does Saving Anticipate Declining Labour Income? An Alternative Test of the Permanent Income Hypothesis" *Econometrica*, 55, 1249-1273.

CAMPBELL, J. and SHILLER, R. (1987). "Cointegration and Tests of Present Value Models" *Journal of Political Economy*, 95, 1062-1088.

COX, D. and HINCKLEY D. (1974). *"Theoretical Statistics"* London, Chapman and Hall.

CURRIE, D. (1981). "Some Long-Run Features of Dynamic Time-Series Models" *Economic Journal*, 363, 704-715.

DAVIDSON, J., HENDRY, D. SRBA, F. and YEO, S. (1978). "Econometric Modelling of the Aggregate Time Series Relationship between Consumers Expenditure and Income in the U.K." *Economic Journal,* 88, 661-632.

DAVIDSON, J. and HENDRY, D. (1981). "Interpreting Evidence: The Behaviour of Consumers 'Expenditure in the U.K." *European Economic Review,* 16, 177-192.

DICKEY, D., BELL, W. and MILLER, R. (1986). "Unit Roots in Time-series Models: Tests and Implications" *American Statistician,* 40, 12-26.

DICKEY, D. and FULLER, W. (1979). "Distribution of the Estimators for Autoregressive Time-Series with a Unit Root". *Journal of the American Statistical Association.* 74, 427-431.

DICKEY, D. and FULLER, W. (1981). "Likelihood Ratio Statistics for Autoregresive Time-Series with a Unit Root" *Econometrica,* 49, 1057-1072.

DICKEY, D. and PANTULA, S. (1987). "Determining the Order of Differencing in Univariate Time-Series" *Journal of Business and Economic Statistics,* 5, 455-462.

DOLADO, J., T. JENKINSON and S. SOSVILLA-RIVERO (1989) "Cointegration and Unit Roots: A Survey" (de próxima aparición en *Journal of Economic Surveys*).

DOLADO, J. BANERJEE, A. and GALBRAITH, J. (1989a). "Recursive Tests for Unit Roots and Structural Breaks in long Annual GNP Series" (de próxima aparición en *Journal of Applied Econometrics*).

DOLADO, J., ERICSSON, N. and J. KREMERS (1989b). "Inference in Conditional Dynamic Models with Integrated Variables" (mimeo).

DOLADO, J. (1988). "Innovación Financiera, Inflación y Estabilidad de la Demanda de ALP en España" *Boletín Económico,* Abril, 19-35.

DHRYMES, P. (1978). *"Mathematics for Econometrics"* New York Springer-Verlag.

EDISON, H. and KLOVELAND, J. (1987). "A Quantitative Reassessment of the Purchasing Power Parity: Evidence from Norway and the U.K." *Journal of Applied Econometrics,* 2. 209-334.

ENGLE, R. HENDRY, D. and J. RICHARD (1983). "Exogeneity" *Econometrica,* 51, 277-304.

ENGLE, R. and GRANGER, C. (1987). "Cointegration and Error Correction: Representation, Estimation and Testing" *Econometrica,* 49, 1057-1072.

ENGLE, R. and YOO, S. (1987). "Forecasting and Testing in Cointegrated Systems" *Journal of Econometrics,* 35, 143-159.

ENGLE, R., GRANGER, C., HYLLEBERG, S. and YOO, S. (1988). "Seasonal Integration and Cointegration" University of S. Diego Discussion Paper 88-32.

ESCRIBANO, A. (1986). "Identification and Modelling of Economic Relationships in a Growing Economy" Ph. D. Dissertation University of California, S. Diego.

ESCRIBANO, A. (1987). "Cointegration, Time Co-trends and Error Corretion Systems: An Alternative Approach" CORE Discussion Paper 8715.

EVANS, G. and SAVIN, N. (1981). "Testing for Unit Roots I'*Econometrica* 49, 753-777.

FRIEDMAN, M. (1957). *"A Theory of the Consumption Function"* Princeton N.J. Princeton University Press.

FULLER, W. (1976). *"Introduction to Statistical Time-Series"*. New York. Wiley.

GOURIEROUX, C., MAUREL, F. and MONFORT, A. (1987). "Regression and Non Stationarity" ENSAE Discussion Paper 8708, Institut National de la Statistique et des Etudes Economiques.

GRANGER, C. and NEWBOLD, P. (1974). "Spurious Regressions in Econometrics" *Journal of Econometrics,* 9, 111-120.

GRANGER, C. (1981). "Some Properties of Time-Series Data and their Use in Econometric Model Specification" *Journal of Econometrics,* 16, 121-130.

GRANGER, C. and WEISS, A. (1983). "Time-series Analysis of Error Correcting Models" in Karlin, S., Amemiya, T. and Goodman, L. (eds). *Studies in Econometric Time-Series and Multivariate Statistics.* New York. Academic Press.

HALL, S. (1986). "An Application of the Granger and Engle Two-Step Estimation Procedure to U.K. Aggregate Wage Data" *Oxford Bulletin of Economics and Statistics,* 48, 229-240.

HENDRY, D. and MIZON, G. (1978). "Serial Correlation as a Convenient Simplification, not a Nuissance: A Comment on a Study of the Demand for Money by the Bank of England" *Economic Journal,* 88, 549-563.

HENDRY, D. and RICHARD, J. F. (1982). "On the Formulation of Empirical Models in Dynamic Econometrics" *Journal of Econometrics,* 20, 3-33.

HENDRY, D. and RICHARD, J. F. (1983). "The Econometric Analysis of Economic Time Series" *International Statistical Review,* 51, 111-163.

HENDRY, D. and ERICSSON, N. (1987). "Assertion Without Empirical Basis: An Econometric Appraisal of 'Monetary Trends... in the U.K.' by M. Friedman and A. Schwartz 'Applied Economics Discussion Paper 25, University of Oxford. (de próxima aparición en American Economic Review).

IMHOF, P. (1961). "Computing the Distribution of Quadratic Forms in Normal Variates" *Biometrika*, 48, 419-426.

JENKINSON, T. (1986). "Testing Neo-Classical Theories of Labour Demand: An Application of Cointegration Techniques" *Oxford Bulletin of Economics and Statistics*, 48, 241-251.

Journal of Economic Dynamics and Control (1988). "Economic Time-Series with Random Walks and other Non Stationary Components" ed. by M. Aoki.

JOHANSEN, S. (1988). "Statistical Analysis of Cointegration Vectors" *Journal of Economic Dynamics and Control*, 12, 231-254.

JOHANSEN, S. and JUSELIUS, K. (1988). "Hypothesis Testing for Cointegration Vectors with an Application to the Demand for Money in Denmark and Finland" Institute of Economics Discussion Paper 88-05. University of Copenhagen.

JEVONS, W. (1884). *"Investigation in Currency and Finance"* London. Mc Millan.

MARAVALL, A. (1979). *"Identificación in Dynamic Shock-Error Models"* New York, Springer-Verlag.

NELSON, C. and PLOSSER, C. (1982). "Trends and Random Walks in Macroeconomic Time-Series" *Journal of Monetary Economics*, 10, 139-162.

OSBORN, D. CHUI, A., SMITH, J. and BIRCHENHALL, C. (1989). "Seasonality and the Order of Integration for Consumption" (de próxima aparición en *Oxford Bulletin of Economics and Statistics*)

Oxford Bulletin of Economics and Statistics (1986). "Economic Modelling with Cointegrated Variables" ed by D. Hendry.

PAGAN, A. and M. WICKENS (1989). "Econometrics: A Survey" (mimeo) (de próxima aparición en *Economic Journal*).

PEÑA, D. and BOX, G. (1987). "Identifying a Simplifying Structure in Time Series" *Journal of the American Statistical Association*, 82, 836-843.

PERRON, P. (1988). "Trends and Random Walks in Macroeconomic Time-Series. Further Evidence from a New Approach" *Journal of Economic Dynamics and Control*, 12, 297-332.

PHILLIPS, A. (1954). "Stabilization Policy in a Closed Economy" *Economic Journal*, 64, 290-323.

PHILLIPS, P. (1986). "Understanding Spurious Regressions In Econometrics" *Journal of Econometrics*, 33, 311-340.

PHILLIPS, P. (1987). "Time Series Regression with a Unit Root" *Econometrica*, 55, 277-301.

PHILLIPS, P. and OULIARIS, S. (1988). "Testing for Cointegration Using Principal Components Methods" *Journal of Economic Dynamics and Control*, 12, 205-230.

PHILLIPS, P. (1988). "Optimal Inference in Cointegrated Systems" Cowles Foundation Discussion Paper 866.

PHILLIPS, P. and PARK, J. (1988). "Asymptotic Equivalence of Ordinary Leasat Squases and Generalized Least Square in Regressions with Integral variables" *Journal of the American Statistical Association*, 111-115.

SALMON, M. (1982). "Error Correction Mechanisms" *Economic Journal*, 92, 615-620.

SARGAN, J. D. (1964). "Wages and Prices in the U.K. A Study in Econometric Methodology" in Hart, P., Mills, G. and Whittaker, J. (e.. *Econometric Analysis for National Economic Planning*, London, Butterworths.

SARGAN, J. D. and BHARGAVA, A. (1983). "Testing Residuals from Least Square Regression for Being Generated by the Gaussian Random Walk. *Econometrica*, 51, 153-174.

SIMS, C., STOCK, J. and WATSON, M. (1986). "Inference in Linear Time-Series Models with Some Unit Roots" (mimeo) (de próxima aparición en Econométrica).

STOCK, J. (1987). "Asymptotic Properties of Least Squares Estimates of Cointegration Vectors" *Econometrica*, 1035-1056.

STOCK, J. (1988). "A Reexamination of Friedman's Consumption Puzzle" *Journal of Business & Economic Statistics*, 6, 401-407.

STOCK, J. and WATSON, M. (1988a). "Variable Trends in Economic Time Series" *Journal of Economic Perspectives*, 2, 147-174.

STOCK, J. and WATSON, M. (1988b). "Testing for Common Trends" *Journal of the American Statistical Association*, 83, 1097-1107.

WHITE, H. (1984). *"Asymptotic Theory for Econometricians"* London, Academic Press.

WORKING, H. (1934). "A Random Difference Series for Use in the Analysis of Time-Series" *Journal of the American Statistical Association*, 2, 11-24.

COINTEGRATION: A GENERAL VIEW

SUMMARY

This paper provides a survey of recent developments in cointegration theory among several integrated time-series. Cointegration implies the existence of equilibrium relationships and therefore offers a generic route to define and test long run relationships among such time-series. The paper reviews results on representation, estimation, testing and prediction in cointegrated systems, addressing also the possibility of cointegration at different time frequencies.

Key words: Cointegration, Trends, Error Correction Mechanism.

AMS 1980: 90A20, 62M20, 62P20.

COMENTARIOS

ALVARO ESCRIBANO SAEZ
Universidad Complutense de Madrid

Quiero comenzar estas líneas dando la enhorabuena a J. J. Dolado por la excelente revisión de la literatura que ha hecho. El tema de cointegración ha sido, y sigue siendo, un tema de gran interés para la profesión ya que sus implicaciones afectan tanto al investigador aplicado como al teórico.

Como toda gran idea, la cointegración cuestiona los fundamentos del estado del arte. En este caso, ha llevado a desarrollar una teoría asintótica distinta de la que se suele encontrar en los libros de texto de Econometría y Estadística Matemática. El cúmulo de nuevos resultados que ello ha acarreado, supone un gran coste de entrada al no experto. El autor ha demostrado a lo largo de los últimos artículos sobre el tema, que tiene una gran capacidad de síntesis y una gran habilidad para seleccionar los aspectos fundamentales de esta literatura y presentarlos de forma asequible para el no especialista.

Para dar una idea de las limitaciones de esta panorámica me centraré en algunos aspectos que no han sido tratados por el autor y que considero de gran importancia. En concreto me voy a referir a cuatro temas diferentes.

El primero se refiere al supuesto de que los procesos estocásticos tienen media cero. Como él mismo reconoce éste es un supuesto simplificador. El problema reside en que es un supuesto que ciertamente no es inocuo. Las implicaciones del mismo son importantes, sobre todo cuando se trata de datos económicos ya que rara vez se cumple, o al menos es un tema debatible. A continuación pasaré a desarrollar más extensamente este punto.

Desde hace muchos años se discute sobre si las series económicas tienen tendencias de tipo determinístico o estocástico. En el fondo esta cuestión se puede extender a si las series económicas tienen tendencias en la media o en la varianza. El suponer que las series tienen media cero nos inclina a pensar que las series tienen tendencias de tipo estocástico y que la tendencia está en la varianza y no en la media, véase Escribano (1987, a). Todo ello se deriva de suponer que las series son integradas con media cero.

Empíricamente la polémica salpica a otras áreas y de ahí que el supuesto sea más fuerte. Por ejemplo, existe en la actualidad una polémica sobre si algunas variables económicas, como son la inversión y el producto interior bruto, son integradas de orden dos, I(2), o son I(1) con tendencias segmentadas en las medias, véase Andrés et al. (1990) para una discusión formal aplicada al caso español.

El supuesto de media cero aún tiene mayores implicaciones cuando se trata del tema de cointegración. Pongamos un ejemplo. Supongamos que dos series son I(1) con tendencias en las medias. Entonces para que se dé una relación de equilibrio a largo plazo, como de la que se habla en el artículo, no es suficiente con que las variables estén cointegradas sino que además es necesario que tengan co-tendencias en la media. De lo contrario, la relación de equilibrio no se observaría en la realidad o se observaría muy poco frecuentemente. Para una discusión de este tema véase Escribano (1987, a).

El segundo tema que quiero resaltar es el siguiente. A la hora de formular el modelo multivariante, el autor parte de la ecuación (6) que es una representación ARIMA (p,q) n-dimensional para de ahí llegar a una representación de Wold multivariante, ecuación (7). Por lo tanto la matriz C(L) de la representación de Wold es una matriz racional de polinomios finitos del operador de retardos, L. El autor podría haber utilizado esta propiedad para mencionar una derivación alternativa de los modelos de corrección de error basada en la descomposición de Smith-McMillan-Yoo, ver por ejemplo Yoo (1986) y Escribano (1990). Esta descomposición ha resultado ser muy útil en la derivación de representaciones alternativas cuando las variables son integradas de órdenes mayores que uno. Una utilización de esta descomposición aplicada a frecuencias estacionales aparecerá próximamente en el número monográfico sobre cointegración de la revista Cuadernos Económicos de ICE.

El tercer tema que quiero resaltar es que cuando el interés reside en la estimación de los coeficientes de la relación del largo plazo, la representación multivariante en forma de mecanismo de corrección de error, ecuación (45), no tiene porqué ser la más apropiada. Bewley (1979) propuso una representación alternativa aunque observacionalmente equivalente, véase ecuación (43). Es sorprendente que el autor no haya mencionado explícitamente las ventajas y desventajas de cada una de ellas sobre todo cuando él mismo cita un artículo suyo que trata específicamente de este tema, véase Banerjee et al. (1988).

El cuarto tema se refiere a la no mención de posibles generalizaciones del concepto de cointegración. Entre ellas está el tema de la multicointegración, los conjuntos de atracción, modelos de corrección de error no lineales, cointegración heterogénea, etc. Los interesados en estos temas pueden consultar, por ejemplo el próximo numero monográfico sobre cointegración de la revista Cuadernos Económicos de ICE y Escribano (1989, b).

Por último, y ya para terminar, quiero agradecer una vez más al autor el esfuerzo simplificador y didáctico que ha realizado para poder llegar a presentar esta excelente panorámica de una forma tan clara y rigurosa.

REFERENCIAS

ANDRES, J., A. ESCRIBANO, C. MOLINAS, D. TAGUAS (1990). Modelización Econométrica con Restricciones de Equilibrio: La Inversión en España. Aceptado para su publicación por Antoni Bosch, Editor.

BANERJEE, A., J. GALBRAITH, J. J. DOLADO (1988). "Dynamic Specification with a General Error Correction Form". (de próxima aparición en Oxford Bulletin of Economics and Statistics).

BEWLEY, R. A. (1979). "The Direct Estimation of the Equilibrium Response in a Linear Model". Economics Letters 3.

ESCRIBANO, A. (1990). "Introducción al Tema de Cointegración y Tendencias". Cuadernos Económicos de I.C.E. Mayo-Junio 1990.

ESCRIBANO, A. (1987, a). "Co-integración, Time Co-trends and Error Correction Systems: An Alternative Approach". CORE. D. P 8715. University of Louvain.

ESCRIBANO, A. (1987, b). "Error Correction Systems: Nonlinear Adjustments to Linear Long Run Relationships". CORE. D. P. 8730. University of Louvain.

YOO, S. (1986). "Multi-cointegrated Time Series and Generalized Error Correction Models". D. P. University of California, San Diego.

F. JAVIER FERNANDEZ MACHO
Departamento de Análisis Económico
e Instituto de Economía Pública
Universidad del País Vasco

Felicito al autor por esta excelente panorámica de los múltiples aspectos que el concepto de cointegración representa en el análisis de series temporales. El esfuerzo dedicado a presentar estos aspectos de manera asequible al lector y la abundancia de ejemplos no puede ser menos loable.

La importancia del concepto de cointegración reside en servir de soporte estadístico para reconciliar la teoría económica, empeñada en sugerir relaciones a largo plazo entre los niveles de ciertas series obviamente no estacionarias, con la práctica econométrica que ha insistido, quizá miópicamente, en el análisis a corto plazo entre transformaciones estacionarias que eviten la obtención de relaciones espúreas. Si a esto añadimos una sencillez tanto conceptual como operativa, resulta ser poco sorprendente que cointegración sea una de esas nociones clave con las que, en esencia, muy pocos pueden mostrarse en franco desacuerdo. Al contrario, cointegración es hoy día un importante catalizador de interesantes trabajos de investigación econométrica.

Es posible, sin embargo, ofrecer algunas consideraciones ampliando o modificando el enfoque, más que la esencia, con que el concepto se maneja habitualmente. Es en este sentido en el que desearía enmarcar el concepto de cointegración en el dominio de la frecuencia. Series cointegradas muestran, en el dominio del tiempo, fluctuaciones comunes a largo plazo (*c. f.* representación TC). Esto se refleja, en el dominio de la frecuencia, en que el espectro de una transformación estacionaria del vector de series exhibe rango incompleto en la frecuencia origen. Obviamente podemos extender la idea anterior a cualesquiera otras frecuencias. En concreto el autor muestra cómo la frecuencia estacional ha recibido recientemente alguna atención, a la que podríamos añadir el modelo multivariante estudiado en Fernández (1986, cap. 7) (*) el cual contiene factores comunes subyacentes con raíces en las frecuencias regular y estacional. Pero además el enfoque espectral permite de esta manera imaginar una interesante extensión del concepto que contemple cointegración en bandas continuas de frecuencias; en este sentido podríamos comprobar, por ejemplo, si ciertos indicadores mensuales del ciclo económico se muestran cointegrados, posiblemente con distintos desfases, en la banda cíclica $(\pi/40, \pi/10)$.

(*) Fernández Macho, F. J. (1986), "Estimation and Testing in Multivariate Time Series Models", Ph. D. Thesis, London School of Economics.

Estas consideraciones son una pequeña muestra de las futuras líneas de investigación en el campo de series cointegradas para las que la presente panorámica sin duda provee una notable introducción.

ANTONIO GARCIA FERRER
Departamento de Análisis Económico
Universidad Autonóma de Madrid

El doctor Dolado nos presenta una panorámica excelente sobre un tema de máxima actualidad en Econometría, como es el de la cointegración. Vaya pues, por delante, mi felicitación al autor por esta detallada exposición sobre un tema en donde sus propias aportaciones recientes validan, si cabe aún más, la brillante exposición realizada en este trabajo. Dado el tremendo interés mostrado por el "nuevo paradigma", y el creciente número de artículos dedicados al mismo, la tarea de escribir esta panorámica sobre el "estado de la cuestión", no es algo sencillo. Por cada día que se pasa sin "cerrar" el artículo, se corre el riesgo de dejar fuera del mismo la última aportación relevante.

No siendo un especialista en cointegración, mis comentarios, necesariamente, van a estar más alejados de los aspectos técnicos del trabajo que de las repercusiones prácticas que esta aproximación puede tener en el trabajo econométrico aplicado de los años venideros. En concreto, hay dos aspectos prácticos de la cointegración que me preocupan: la implicación teórica de la supuesta presencia de raíces unitarias en muchas series económicas; y la estrategia de construcción de modelos econométricos sobre la base de variables cointegradas. Sobre estas dos cuestiones centraré mis comentarios.

Antes de entrar de lleno sobre estos aspectos quisiera, sin embargo, comentar una cuestión previa sobre la hipótesis de estacionariedad en el trabajo econométrico aplicado con anterioridad a la aparición del concepto de cointegración. En la Introducción del trabajo se indica que dicho supuesto apenas se cuestionaba y que el análisis económico procedía como si todas las series económicas fuesen estacionarias. Esta afirmación me parece algo exagerada y no hace justicia al trabajo pionero de Box y Jenkins hace ya más de veinte años [Por cierto, que el nombre de ARIMA lleva implícito el concepto procesos "integrados"]. Desde entonces —con mayor

(*) Trabajo a publicarse en la Revista Estadística Española.

o menor retraso— se ha ido reconociendo que las técnicas usuales del análisis de regresión suelen proporcionar resultados muy insatisfactorios en aquellas situaciones donde las variables contienen tendencias estocásticas. El influyente trabajo posterior de Granger y Newbold proporcionó (mediante un sencillo ejemplo) una demostración empírica de los peligros reales de la utilización de modelos estáticos en econometría y de la posibilidad de generar regresiones espurias con series temporales no estacionarias. En este mismo sentido, quisiera mostrar mi desacuerdo con el doctor Dolado sobre la interpretación de las funciones de transferencia y su aparente "incapacidad para detectar las propiedades a largo plazo del modelo ignorando, por tanto, las relaciones de equilibrio sugeridas por la teoría económica" (pág. 3). Dentro del campo de las series temporales existen numerosos trabajos [véase, inter alia, del Hoyo y Terceiro (1983), García-Ferrer y del Hoyo (1987) y Flores (1987)] en donde utilizando esta aproximación se analizan las implicaciones a corto y largo plazo de los distintos modelos obtenidos.

Por lo que se refiere a la "epidemia" de estudios sobre raíces unitarias en economía, esta podría ser el resultado de dos tipos de ideas que fluyen dentro del panorama actual. Por un lado, está la cuestión sobre si la existencia de series económicas con raíces unitarias está íntimamente ligada o no con cuestiones importantes acerca del comportamiento económico de los agentes. Por otro, existe una sensación generalizada de que la presencia de raíces unitarias en los datos complica considerablemente el análisis estadístico y hace necesarios nuevos procedimientos válidos para este caso. La lista de ejemplos típicos que se adscriben a alguna de estas dos ideas es exhaustiva: desde la hipótesis de mercados eficientes a los tipos de cambio, y desde la función de consumo a la modelización univariante del PNB. Sin embargo, también empiezan a ser numerosos los autores [véase, inter alia, Sims (1988) y Christiano (1989)] que ponen seriamente en duda la relevancia teórica de esta aproximación e insisten en que las implicaciones teóricas de la presencia de raíces unitarias sólo son válidas bajo determinadas hipótesis especiales (1). Asimismo, y desde un punto de vista empírico, es posible demostrar la variabilidad considerable en los parámetros de los modelos de "paseo aleatorio" cuando éstos se estiman a través de un algoritmo recursivo [del Hoyo y García Ferrer (1989)].

(1) Por ejemplo, Sims (1984) demuestra como la intuición detrás de la hipótesis de mercados eficientes es aplicable realmente como una aproximación en pequeños intervalos de tiempo.

Directamente relacionados con la cuestión anterior aparecen los tests de cointegración como tests sobre la existencia de raíces unitarias. Tomemos el Ejemplo 2 (p. 12) del trabajo en donde el test sobre cointegración es en realidad un test condicional (que depende de los valores de ρ_1 y ρ_2) a que z_t e y_t sean procesos I(1). Como en realidad los valores de ρ_1 y ρ_2 son *desconocidos*, la potencia de los tests de raíces unitarias contra sus alternativas es, a menudo, muy baja y la posibilidad de cometer errores tipo II es considerable. De hecho, una *limitación muy importante* de los tests de cointegración es que —dados los tamaños típicos de las series económicas— poseen un nivel de discriminación muy bajo a la hora de elegir entre distintas hipótesis, especialmente en el caso de más de dos variables. Por ejemplo, para tamaños muestrales comprendidos entre 100 y 150 observaciones es difícil distinguir entre un proceso integrado y otro estacionario pero serialmente correlacionado. En este tipo de situaciones la aplicación de los tests bayesianos usuales o del recientemente propuesto por Sims (1989) parece altamente aconsejable. En este sentido, la batería de tests clásicos aplicados a la cointegración puede plantear serios problemas en aquellos casos donde existan situaciones de conflicto y la decisión final tenga que basarse sobre consideraciones asintóticas (2). Por otra parte, y dado el interés que tiene la modelización econométrica con datos de periodicidad distinta a la anual, es importante hacer constar que, con excepción del caso trimestral. [Hylleberg et al. (1988)] no se han desarrollado todavía los tests de cointegración que permitan aplicar esta aproximación a datos mensuales o semanales no estacionarios (3).

Finalmente, me gustaría comentar brevemente qué tipos de estrategias de construcción de modelos econométricos podrían resultar al aplicar el concepto de cointegración. Por estrategia, entiendo todo el proceso de construcción de un modelo: desde la etapa de especificación hasta la de predicción post-muestral. En este sentido, y dada la relativa escasez de trabajos empíricos que conozco en esta línea, la discusión ha de ser necesariamente especulativa. Sin embargo, como siempre que aparecen aplicaciones sobre una nueva metodología, me preocupa que una utilización poco cuidadosa de la misma pueda llevar a determinadas conclusiones sobre política económica cuya comprensión resulta difícil. Es más, ya he oído públicamente a más de un especialista congratularse de la aparición del concepto de

(2) Un reciente trabajo de Schwert (1989) utilizando experimentos de Monte Carlo demuestra como los tests sobre raíces unitarias propuestos Said y Dickey, Phillips y Phillips y Perron, poseen distribuciones finitas distintas que los tradicionales de Fuller y Dickey-Fuller para procesos AR.

(3) Véanse, por ejemplo, las conclusiones de Baillie y Selover (1987) a la hora de justificar el fracaso del modelo monetario de los tipos de cambio.

cointegración porque "demuestra que la estimación por MCO de relaciones estáticas en niveles es básicamente correcta". También he visto como se "caían" determinadas variables de una "relación de equilibrio" sobre la base de tests asintóticos con muestras de 20 observaciones. Yo mismo [García Ferrer (1988)] he encontrado resultados sorprendentes sobre la ausencia de cointegración entre consumo y renta para el caso español entre 1954 y 1984, o entre fecundidad y nupcialidad en España a lo largo del presente siglo. ¿Es razonable afirmar que este tipo de relaciones lógicas no constituyen una relación de equilibrio? ¿Qué estrategia seguir cuando las series analizadas presentan valores atípicos importantes como es el caso de las series demográficas o del propio PNB español?

Mi comentario final tiene, por tanto, que referirse a un deseo de ver en funcionamiento práctico las ideas expresadas en esta panorámica, y comprobar si el supuesto potencial teórico que el nuevo concepto arrastra se ve confirmado por un conjunto de aplicaciones más convincentes que las realizadas hasta ahora. Ello no obsta, para que nuevamente felicite al doctor Dolado por el espléndido trabajo que nos ha presentado.

REFERENCIAS

BAILLIE, R. T. y D. D. SELOVER (1987). "Cointegration and Models of Exchange Rate Determination". *International Journal of Forecasting,* 3, 43-51.

CHRISTIANO, L. (1989). "On Unit Roots in GNP: Do we know and do we care? *NSF-NBER Time Series Seminar,* Madrid, 11/12 Septiembre.

FLORES, R. (1987). "Análisis Econométrico sobre la Incidencia de la Economía de Estados Unidos y la Economía Mundial en la Economía Española". Tesis Doctoral no publicada, Departamento de Econometría, Universidad Complutense, Madrid.

GARCIA FERRER, A. (1988). Comentarios al trabajo de J. L. Raymond y E. Uriel "Investigación Econométrica Aplicada: Un Caso de Estudio". *I Seminario sobre Especificación y Validación de Modelos Econométricos.* Zaragoza, mayo.

GARCIA FERRER, A. y J. del HOYO (1987). "Analysis of the Car Accident Indexes in Spain: A Multiple Time Series Approach". *Journal of Business and Economic Statistics,* 5, 27-38.

DEL HOYO, J. y A. GARCIA FERRER (1989). "Sobre la Constancia de los Parámetros en los Modelos de Paseo Aleatorio del PNB". Documento de trabajo WP # 7/89, Departamento de Análisis Económico, Universidad Autónoma de Madrid.

DEL HOYO, J. y J. TERCEIRO (1983). "On the Interpretation of Long-Run and Short-Run Effects in Transfer Function Models. Documento de Trabajo WP # 3/83, Departamento de Econometría, Universidad Autónoma de Madrid.

HYLLEBERG, S. R. ENGLE, C. W. J. GRANGER y B. S. YOO (1988). "Seasonal Integration and Cointegration". Discussion Paper # 88-32. Department of Economics. University of California, San Diego.

SCHWERT, G. W. (1989). "Tests for Unit Roots: A Monte Carlo Investigation". *Journal of Business and Economic Statistics,* 7, 147-159.

SIMS, C. A. (1984). "Martingale-like Behavior of Prices and Interest Rates", Center for Economic Research Discussión Paper N.º 205 (University of Minnesota, Mineapolis, MN).

SIMS, C. A. (1988). "Bayesian Skepticism on Unit Roots Econometrics". *Journal of Economic Dynamics and Control,* 12, 463-474.

CESAR MOLINAS
Ministerio de Economía y Hacienda

El trabajo de Dolado es, en mi opinión, una exposición muy buena de las motivaciones y de los desarrollos recientes de la teoría de variables cointegradas. De hecho es la más completa, rigurosa y —a la vez— intuitiva que conozco y, por todo ello, creo que hay que celebrar que esté escrita en castellano.

En el marco conceptual de la teoría de la cointegración se supera la antigua dicotomia entre estimación en niveles versus estimación en tasas de crecimiento. Además queda realzado el papel de la Teoría Económica en la especificación de relaciones econométricas, al quedar mejor delimitado el largo plazo de las mismas. Por estas y otras razones, la teoría de la cointegración proporciona una disciplina intelectual de gran utilidad para el economista aplicado.

Sin embargo sus aplicaciones no están exentas de problemas serios. Voy a referirme, brevemente, a dos de ellos. En primer lugar, el orden de integrabilidad de muchas series macroeconómicas parece ser más una propiedad de las muestras que de las series en sí. Es difícil, por ejemplo, rechazar que el PIB real español sea una variable integrada de orden 2 con datos de la Contabilidad Nacional desde 1964. Lo mismo ocurre con algunas componentes de la demanda agregada. Cuesta, sin embargo, creer que en períodos más prolongados de tiempo este tipo de variables exhiba una doble raíz unitaria. Ello implicaría admitir un grado de variabilidad que parece poco razonable desde una perspectiva más secular. En cualquier caso, determinadas variables pueden aparecer como cointegradas en una muestra y como no cointegradas en otra debido a este tipo de problemas, y ello puede comprometer la buena especificación de según qué modelos econométricos.

En segundo lugar está el problema de cómo contrastar la existencia o no de cointegración entre diversas variables. En otro sitio (ver Molinas (1986)) he mostrado que la mayoría de los contrastes de cointegración disponibles para el economista aplicado, son extremadamente sensibles a la hipótesis de que los errores de la relación estática de cointegración sigan procesos autorregresivos. Si esto no se cumple, los errores de Tipo I, es decir el rechazo de la hipótesis de no cointegración cuando es verdadera, pueden producirse con una frecuencia mucho mayor que la que da el nivel de significación del contraste. Ello es particularmente cierto cuando las variables que se están investigando siguen procesos integrados de media móvil, hecho frecuente en series económicas. El peligro de aceptar como relaciones de cointegración regresiones completamente espurias es, en ese caso, grande, El "caso de Molinas", como amablemente lo bautizó Dolado, puede muy bien ser el de muchos investigadores aplicados.

REFERENCIAS

MOLINAS, C. (1986): "A Note on Spurious Regressions with Integrated Moving Average Errors" Oxford Bulletin of Economics and Statistics, vol. 48, n.º 3.

ALFONSO NOVALES
Universidad Complutense de Madrid y FEDEA

Una vez más, hemos de agradecer al Comité de Redacción de Estadística Española el haber seleccionado uno de los temas más debatidos en la práctica econométrica actual como elemento de discusión en las páginas de la revista. Personalmente, agradezco la invitación recibida para escribir unos comentarios acerca del trabajo de J. J. Dolado, uno de los contrasta-dos especialistas en la materia. Su artículo representa un examen bastante completo, si bien no elemental, de los resultados estadísticos concernientes al tema de la cointegración de variables en un modelo lineal.

No hay mucho que pueda decir acerca de la teoría estadística que se ha desarrollado en este área, pues soy relativamente ajeno a dicha línea de investigación. Como economista empírico, sigo los desarrollos que en ella se vienen produciendo, y quisiera simplemente exponer en estas páginas mi valoración acerca de la relevancia práctica del problema, por si ello pudiera servir de elemento de discusión.

LA EXISTENCIA DE RAICES UNITARIAS EN TEORIA ECONOMICA

Una de las primeras cuestiones que conviene aclarar es la excesiva fijación por las implicaciones que los modelos de equilibrio económico puedan tener acerca de la existencia de raices unitarias en variables económicas. En muchas ocasiones, dicha impresión ha sido producida por una lectura inapropiada de dichos modelos.

Creo que un caso sintomático lo representa el comportamiento del consumo agregado. A partir del modelo de Hall (1978), comenzaron a oirse (y leerse) afirmaciones similares a ésta: "Bajo expectativas racionales, si los agentes optimizan una función de utilidad intertemporal, el consumo obedece un proceso de camino aleatorio" (ver una referencia no explícita en la introducción del artículo de J. J. Dolado). De este modo, tras dicho trabajo, apareció en la literatura la tradicional secuencia de artículos tratando de contrastar dicha hipótesis utilizando datos de los más diversos países. Y, sin embargo, todo lo que un modelo como el descrito asegura es que, en equilibrio, se tiene:

$$(1) \qquad E_t\, U_{c_{t+1}} = \frac{1 + r_t}{1 + \delta} \;.\; E_t\, U_{c_t}$$

donde U_{c_t} denota la utilidad marginal del consumo y δ es el factor de descuento intertemporal. Esta condición de equilibrio no afirma mucho acerca del comportamiento del consumo en su modelización univariante, por dos razones (Véase Novales (1989)): a) porque es una condición sobre la utilidad marginal del consumo, no sobre éste, y b) porque es una condición no-lineal entre los tipos de interés y la utilidad marginal.

Para poder derivar de (1) una condición sobre el consumo, es preciso suponer el tipo de interés real r_t constante. Si además se supone igual a δ, y si la función de utilidad se supone separable en el tiempo (y logarítmica, por ejemplo), se tendrá la raíz unitaria en la serie de consumo como una implicación teórica. Este conjunto de supuestos es suficientemente restrictivo como para relativizar en gran medida el tipo de análisis empírico antes descrito.

Por otra parte, la citada observación es una característica común a los modelos de equilibrio intertemporal bajo expectativas racionales, como muestran las versiones lineal-cuadráticas de Sargent (1979), donde la variable endógena para la que se resuelve el modelo como función de las "forcing variables" tiene un coeficiente en su propio retardo estrictamente inferior a la unidad [1].

Dentro de los modelos financieros, la tantas veces mencionada raíz unitaria de las series de tipos de interés y tipos de cambio, no es estrictamente válida desde el punto de vista teórico sino bajo supuestos igualmente restrictivos (ausencia de costes de transacción y neutralidad al riesgo, entre otros).

Lo que me sugieren estas consideraciones es que, si bien es cierto que la teoría económica debe ser el soporte fundamental de todo trabajo empírico, no es menos cierto que existe un buen número de cuestiones que son de total relevancia para un análisis empírico riguroso, y acerca de los cuales la Teoría Económica no puede manifestarse. Uno de tales aspectos es la modificación que experimentan las distintas implicaciones teóricas al variar la frecuencia de observación de los datos. Podría conjeturarse que si una proposición teórica resulta validada en un análisis empírico, puede encontrarse una frecuencia diferente de observación de las variables para la que dicha proposición será rechazada. Por ejemplo, la implicación de paseo aleatorio para precios financieros debe entenderse en lo que vale, pero no debe interpretarse como un salvaconducto para justificar su contrastación empírica con todo tipo de datos y frecuencias de observación. Dicho de otro modo, el posible rechazo de tal hipótesis con datos mensuales (construidos como medias de datos diarios) del índice de la bolsa de Madrid, no alteraría mi interpretación y valoración de la hipótesis de paseo aleatorio para los precios de activos en mercados eficientes.

---

1 El que la variable endógena obedezca un proceso AR(1) al que se incorporan las expectativas de valores futuros de variables exógenas, se debe a que la especificación de la función objetivo no incluye, generalmente, más de un retardo, por ejemplo, en los términos de coste de ajuste del stock de capital.

CONTRASTES DE COINTEGRACION

La segunda dificultad que encuentro en esta literatura surge de la debilidad (baja potencia) que los contrastes de cointegración tienen en la práctica. Ello hace muchas veces difícil de rechazar la hipótesis de cointegración, aunque tampoco se rechazarían otras alternativas. En ocasiones, ello conduce al investigador a examinar cuidadosamente los residuos de una determinada regresión para tratar de detectar indicios de no estacionariedad y, en caso contrario, no rechazar la hipótesis de que los coeficientes estimados constituyen el vector de cointegración (véase el final de la sección 4.3 en el artículo de Dolado).

Esta práctica es la que me resulta realmente interesante, pues considero que es la única válida en muestras del tamaño que habitualmente utilizamos en el trabajo empírico. El análisis riguroso de los residuos de cualquier modelo econométrico es una práctica esencial y la mejor lección que supuestamente hemos aprendido de los últimos 20 años de análisis de series temporales. Es desde este punto de vista que contemplo la literatura sobre cointegración: es, sin duda, una llamada de atención fundamental para todo analista de series temporales. Es, bastante más que esto, para todo usuario de las distintas variantes de los métodos de estimación minimocuadráticos en modelos de regresión (veánse las recomendaciones en Stock y Watson (1988)).

Una dificultad adicional estriba en el hecho de que las longitudes muestrales con que en ocasiones nos vemos obligados a trabajar no son excesivas, y dificultan en gran manera la detección de relaciones entre variables que sólo se manifiestan a través de un ciclo económico completo. Ello no es un problema fundamental en un análisis que trate de detectar la estructura dinámica de las respuestas a corto plazo entre variables, pero sí puede serlo en lo que afecta al análisis de cointegración. Bien podría ocurrir que dos agregados macroeconómicos no aparezcan como cointegrados de acuerdo con los tests habituales (véanse los comentarios de A. García Ferrer en este mismo volumen para un ejemplo), y difícilmente debe negarse la existencia de relación de equilibrio a largo plazo entre dichas variables.

En definitiva, y en contra de las tajantes afirmaciones que en ocasiones se efectúan (véase la introducción al artículo de Dolado así como los comentarios en la Sección 4.3) la detección mediante procedimientos de estadística clásica de vectores de cointegración no debe considerarse condición necesaria ni suficiente para creer que el vector de variables en estudio satisface una relación de equilibrio a largo plazo. Del mismo modo que tampoco es cierto que otras metodologías alternativas de datos económicos no puedan caracterizar las relaciones de equilibrio a largo plazo entre variables.

Pero hay otros dos aspectos aún por desarrollar en esta línea de investigación que, unidos a los ya citados, hacen que considere aún con bastante precaución sus conclusiones: uno es el relativo al tratamiento de la habitual ausencia de estacionariedad estacional de las variables económicas, observadas mensualmente o con mayor frecuencia. El segundo aspecto es el de la dificultad de obtener la distribución asintótica de modo explícito en muchos casos, lo que conduce a la necesidad de construir mediante procedimientos de Monte Carlo tablas estadísticas que son únicamente válidas para el caso particular para el cual se construyeron. En este sentido, el análisis no paramétrico, basado en procedimientos del tipo "bootstrapping" podría resultar más interesante que las prácticas citadas.

Para finalizar, ha aparecido recientemente una propuesta bien diferente (Sims (1989)) acerca de la literatura de raices unitarias que considero que debiera ser de gran estímulo para quienes nos dedicamos al análisis económico. En ella, se recupera la tradición bayesiana más pura para sugerir (por mucho que resulta imposible resumir el trabajo en un par de afirmaciones) que la detección de raíces unitarias puede deberse en muchos casos a la ausencia de formalización explícita de la opinión del investigador acerca de los valores de los parámetros objeto de interés; en un sentido más preciso, Sims ilustra con varios ejemplos la validez del tratamiento bayesiano y lo inapropiado de las distribuciones a priori "planas" en situaciones próximas a la existencia de raices unitarias. Lo interesante es que creo que muchos estaríamos dispuestos a considerar que esta es la situación habitual en las relaciones entre variables ecónomicas.

Creo que el progreso en una línea de investigación como la que discutimos en este número de Revista Española debe suponer el tratar de encontrar a las dificultades citadas, pero también el tener respuesta a alternativas como la que acabo de citar. El artículo de J. J. Dolado debe verse por todos quienes analizamos series de datos económicos como una contribución muy relevante en cuanto que trata de presentar el estado de una cuestión que impregna todo trabajo empírico y que se halla en estado de plena ebullición. Todo ello hace que el trabajo de Dolado haya sido seguramente de ardua realización. Sus colegas le debemos estar agradecidos.

BIBLIOGRAFIA

STOCK, J. y M. WATSON (1988) "Variable trends in economic time series", *Journal of Economic Perspectives,* vol. 2, 3, pp. 147-174.

SARGENT, T. (1979) *Macroeconomic Theory,* Academic Press, New York.

NOVALES, A. (1989) "The role of adjustement costs in the determination of interest rates", manuscrito.

SIMS, C.A. (1989) "Bayesian skepticism on unit root econometrics", *Journal of Economic Dynamics and Control.*

J. IGNACIO PEÑA

Departamento Análisis Económico

Universidad Autónoma de Madrid

En primer lugar, quisiera felicitar al doctor Dolado por su interesante trabajo. Es un detallado repaso a las principales ideas de un área que muchas personas en la profesión consideran como el nexo de unión entre los enfoques de la Econometría "tradicional" y "de series temporales", y a la que el propio autor ha hecho aportaciones relevantes.

Mis comentarios se van a referir al apartado del trabajo relacionado con los contrastes de cointegración, desde el punto de vista de los contrastes de raíz unitaria.

El contraste más recomendado en la literatura y más empleado en aplicaciones empíricas es el de Dickey-Fuller extendido (EDF), en su formulación por Said y Dickey (1984). Esencialmente consiste en una regresión de la serie diferenciada sobre la serie en niveles retardada un periodo y sobre un conjunto de retardos de la propia serie en diferencias. El número de retardos se escoge de tal forma que la serie residual sea ruido blanco. El contraste de la *t* asociado al coeficiente de la serie en niveles sigue asintóticamente una distribución tabulada en Fuller (1976).

Sin embargo, aunque supuestamente válido para cualquier proceso ARMA (p,q), en un reciente trabajo Schwert (1989) muestra unos extensos ejercicios de simulación para modelos ARIMA (p,d,q) $(p \leq 1,\ d \leq 1,\ q \leq 1)$ en los que se aprecia claramente cómo las distribuciones empíricas de los contrastes EDF pueden ser muy diferentes de las distribuciones de referencia, si el valor del parámetro de media móvil es elevado (superior a 0.5 en general), incluso con tamaños muestrales de 1.000 datos. Además, en el mismo artículo, se expone cómo los contrastes aparentemente más generales de Philips (1987) y Philips y Perron (1989) no sólo no tienen un comportamiento más fiable, sino que si los valores del parámetro de media móvil son superiores a 0.7, los contrastes no se distribuyen de forma parecida a su distribución asintótica ni siquiera para muestras de tamaño 10.000.

Una vía alternativa podría ser el contraste de Solo (1984) basado en el Multiplicador de Lagrange. Para modelos ARIMA $(p,1,q)$ se formula como sigue: dado el modelo $(t = 1, ..., T)$.

$$(1-a(1)B)X(t) - a(2)DX(t-1) - ... - a(p)DX(t-p) = c(B)\ e(t) \qquad (1)$$

$$DX(t) = X(t) - X(t-1)$$

$$c(B) = 1 + c(1)B + ... + c(q)B^q$$

estimar el modelo (1) bajo la hipótesis nula $(a(1) = 1)$ y obtener los residuos estimados $e'(t)$. Construir los regresores $n(t)$ como

$$n(t-1) = X(t-1) \,/\, \hat{c}(B)$$

y efectuar la regresión de $e'(t-1)$ sobre $n(t-1)$, obteniendo el contraste LM como el TR^2 de la anterior regresión. Este contraste no se distribuye chi-cuadrado (ya que la serie $X(t)$ no tiene varianza finita) sino que sigue las mismas distribuciones tabuladas por Fuller (1976). La potencia de éste contraste en muestras finitas aún no se ha estudiado, aunque asintóticamente es equivalente al EDF.

Desde el punto de vista de su aplicación, el problema básico de los contrastes EDF y sus derivados proviene de la no linealidad de la parte MA, ya que las aproximaciones autorregresivas hacen que la potencia del contraste sea baja. Sin embargo si adoptamos el enfoque de Solo (1984) el problema se traslada a la correcta selección de los órdenes p y q del modelo ARMA en cuestión.

El análisis de las posibles consecuencias de éstos hechos presenta una interesante vía de investigación que trabajos como el del Dr. Dolado seguramente estimularán.

REFERENCIAS

FULLER, W. (1976) *The Statistical Analysis of Time Series.* Wiley.

PHILIPS, P. C. B. (1987) Time Series Regression with Unit Roots. *Econometrica,* 55, 277-301.

PHILIPS, P. C. B. y PERRON, P. (1989) Testing for a Unit Root in Time Series Regression. *Biometrika.*

SAID, S. E. y DICKEY, D. (1984) Testing for Unit Roots In ARMA models of Unknown Order. Biometrika, 71, 599-607.

SCHWERT, G. (1989) Tests for Unit Roots: A Monte Carlo Investigation. *Jour. of Bus. and Econ. Stat.,* 7, 147-160.

SOLO, V. (1984) The Order of differencing in ARIMA models. *Jour. of the Amer. Stat. Assoc.,* 79, 916-921.

ARTHUR B. TREADWAY
Departamento de Economía Cuantitativa
Facultad de Ciencias Económicas y Empresariales
Universidad Complutense de Madrid

Box y Jenkins publicaron su obra fundamental del moderno Análisis de Series Temporales (AST) en 1970, aunque todavía hay muchos autores que lo desconocen. La única mención que recibe el enfoque B-J en esta panorámica (tercera frase del cuarto párrafo), ofrece una caracterización errónea del enfoque. El enfoque B-J no se limita al estudio de las variables diferenciadas.

Que la típica serie temporal económica no es estacionaria no fue descubierto hace poco tiempo. El analista B-J descubre esta importante regularidad tan pronto como analice series económicas. Esto me ocurrió hace 13 años; les habría ocurrido a Box y a Jenkins hace más de 25. El reconocimiento de esta regularidad no requiere el empleo de ninguno de los contrastes formales "de raices unitarias" como los inventados por Dickey y otros. De hecho, la aplicación exclusiva de tales contrastes sin emplear los otros métodos B-J apropiados, constituye una práctica poco recomendable aunque muy extendida entre autores de Econometría.

La Cointegración (CI), que entiendo en la versión general no solamente contemporánea, es un tipo de restricción del modelo multivariante estocástico (MS) general. El analista B-J de series temporales sabe detectar, contrastar e imponer esta clase de restricción en muchos tipos de MS sin referirse a la literatura que el Dr. Dolado resume y sin emplear la llamada Teorema de Representación de Granger (TRG) o el llamado Modelo de Corrección de Errores (MCE), dos desarrollos que me parecen muy poco relevantes.

Son muchas las veces que he visto contrastada la hipótesis que ahora se bautiza con el nombre de Cointegración. Se ha rechazado en casi todos los casos que conozco. Mi experiencia indica que la CI no es un fenómeno frecuente en contextos económicamente interesantes.

Cuando el analista B-J elabora un modelo de transferencia (UT), comienza realizando un análisis univariante estocástico (US) de cada variable output e input que ha elegido. Después de formular un UT conceptual inicial, realiza las operaciones de identificación (especificación empírica) de las funciones de transferencia que describen la dependencia del output respecto a los inputs, y pasa a la a estimación eficiente del UT inicial, empleando una estructura ARIMA del ruido igual a la estructura US de la variable output. Así comienza el proceso iterativo de especificación-estimación eficiente-diagnosis-reformulación, proceso que termina cuando el analista ya no encuentra evidencia de mala especificación. Si el número de diferencias en la estructura del ruido al final del proceso es menor que el del modelo US de la variable output, entonces el analista ha detectado la CI; si estos dos números de diferencias no difieren, el analista no ha encontrado la CI. Es típico que las variables presenten estructuras IMA, es decir, en los modelos US alguna(s) diferencia(s) se verá(n) parcialmente compensada(s) por estructuras MA(1) invertibles con parámetro(s) positivo(s). Entonces la CI ocurre en el UT cuando un factor MA(1), presente e invertible en el US del output, resulta no-invertible en el ruido del UT. No se encuentra la no-invertibilidad relevante en este contexto con frecuencia.

Me resulta llamativo que el Dr. Dolado y sus autores favoritos profesan creer que el mundo observado por la Econometría combina la no-estacionariedad de la típica serie temporal económica y la CI como fenómeno frecuente. Esta combinación implica que el económetra no omite de su modelo ninguna variable económica típica en la(s) llamada(s) relación(es) de CI. Así se podría definir una especie de paradiso del económetra, un estado libre de ilusión que no me parece sensato esperar alcanzar salvo como excepción accidental.

El Dr. Dolado sugiere que la CI es implicación de muchas teorías económicas plausibles. Pero, cuando ocurre la CI, no puede existir ninguna variable omitida respecto a la cual la teoría económica en cuestión necesitaría una cláusula *ceteris paribus* en su formulación de la relación de CI. Son muy pocas las teorías económicas plausibles que se pueden expresar sin ninguna cláusula *ceteris paribus.* No creo que exista nada especial en la Teoría Económica que favorezca la CI.

La restricción CI del MS (6) es una restricción de los parámetros autorregresivos exclusivamente; no restringe ni los parámetros media-móvil ni la matriz de varianzas-covarianzas contemporáneas. La CI no resuelve, ni pretende resolver, el problema clásico de identificación econométrica. No ofrece una interpretación económica de las correlaciones cruzadas contemporáneas. Esto implica que ninguna de las representaciones que se maneja en esta literatura (las del llamado TRG, incluyendo el MCE) especifica ecuaciones de comportamiento, es decir, todas ellas son vacías desde la perspectiva económica.

He leído los trabajos que el Dr. Dolado cita (en el primer párrafo después del Ejemplo 1) como ejemplos de teorías económicas que sugieren la CI. Con pocas excepciones, que no incluyen los trabajos sobre la economía española, ni las teorías económicas citadas ni los informes de los análisis de datos me resultan convincentes. Me parece que se impone la hipótesis de CI en muchos de estos trabajos cuando la evidencia empírica adecuadamente evaluada no la apoya.

Contestación

JUAN J. DOLADO

Los comentarios de mis distinguidos colegas son muy interesantes y constructivos y, sin duda alguna, enriquecen y sitúan en una perspectiva más amplia algunas de las ideas vertidas en mi artículo. Vaya por delante mi profundo agradecimiento a todos ellos. Antes de proceder con aspectos concretos de la contestación, quisiera, sin embargo, tomar prestadas las siguientes palabras de Antonio García Ferrer:"... Por cada día que se pasa sin "cerrar" el artículo, se corre el riesgo de dejar fuera del mismo la última aportación relevante". Sirva tan amable frase para disculparme por el hecho de que muchos de los puntos que contienen los comentarios constituyen, en alguna medida, aportaciones relevantes que, a la hora de entregar el artículo a los editores, a principios de 1989, todavía estaban en su proceso de génesis. Creo que es un hecho característico que la literatura sobre raíces unitarias y cointegración, tanto en sus aspectos económicos como estadísticos, tiene un componente tendencial, a la vista del gran número de artículos que los editores de las revistas profesionales especializadas están incluyendo en sus últimos números. Con ello quisiera en parte disculparme por la posible omisión de aspectos de interés en la panorámica y agradecer de nuevo la tarea de los comentaristas en atraer la atención del lector sobre los mismos.

Siguiendo mis preferencias de una metodología que vaya de "lo general a lo particular" articularé mi contestación en tres apartados con grado decreciente de generalidad, tratando de cubrir la mayoría de las cuestiones contenidas en los comentarios.

1. LA IMPORTANCIA DEL CONCEPTO DE COINTEGRACION

El profesor Treadway descalifica el enfoque, caracterizándolo como un fenómeno poco frecuente en contextos económicamente interesantes. Me resulta difícil entender un mundo en que cada proceso estocástico posea una raíz unitaria y no exista ninguna cointegración. Con ello no pretendo dogmatizar sobre la generalidad del concepto sino establecer un marco analítico en que sea posible su contrastación. Personalmente cuando trato de construir un modelo procuro comenzar por formular una teoría sobre las relaciones de cointegración potencialmente existentes entre las variables elegidas. Por ejemplo si se trata de analizar la relación entre consumo y renta me inclino a pensar que se hallan cointegrados y, en aspectos específicos, la teoría económica me ayuda a elegir entre diversas medidas de las variables relevantes (renta disponible y no renta del trabajo, de acuerdo con Campbell (1987)). Por supuesto no restringiría la velocidad ajuste a la solución de equilibrio ya que ésta varía entre diferentes países y depende posiblemente de los mecanismos financieros existentes. A menudo uno diseña el modelo en función de las posibles restricciones que los agentes puedan experimentar en diversos mercados, etc. (véase p. ej., el excelente trabajo de Andrés et al. (1990) sobre el tema). En este sentido creo que la teoría económica tiene algo que aportar y comenzar por imponer todas las raíces unitarias en el sistema, me parece poco interesante. Estoy de acuerdo que desde otras, aparentemente distintas, metodologías como la descrita por el profesor Treadway en su exposición del enfoque UT, podría efectuarse el mismo tipo de análisis, si bien en la mayoría de los casos la práctica habitual no suele ser tan rigurosa como el propio Treadway. En cualquier caso el enfoque no está exento de tantas o más dificultades como las encontradas en los métodos descritos en el artículo, i.e. el llamado "pile up problem" en la estimación de raíces unitarias en la media móvil, la interpretación estructural del modelo y su tendencia a largo plazo, etc. En cualquier caso, para su satisfacción, deseo manifestarle que una interpretación mucho menos dogmática y más rigurosa y constructiva de su enfoque también puede encontrarse en la literatura "cointegradora" (véase Phillips y Ouliaris (1988) donde se adopta la existencia de cointegración como hipótesis nula) sin que parezcan existir motivos claros para decantarse por la misma.

2. LA POTENCIA DE LOS CONTRASTES DE COINTEGRACION

Los profesores García Ferrer, Molinas, Novales y Peña muestran su preocupación por la baja potencia de los contrastes de integración y cointegración con los tamaños muestrales con los que habitualmente trabajamos los econόmetras. Supongo que esto responde a una ley de vida por la que distinguir entre una raíz de 0.95 y una raíz de 1.0 con una muestra finita resulta poco menos que imposible. Existen, sin embargo, diversos enfoques que mejoran las propiedades de los contrastes, muchos de ellos citados por los propios autores en sus comentarios. A modo de ejemplo, citaré el trabajo de Hall (1989) en el marco de modelos con medias móviles altas, estimando por variables instrumentales, o el procedimiento diseñado por Hansen (1990) basado en la iteración del método Cochrane-Orcutt. Javier Hidalgo y yo mismo (1990) hemos intentado refinar el procedimiento de Said y Dickey (1985) comentado por el profesor Peña. En cualquier caso, como bien ha puesto de manifiesto Phillips (1988), la teoría asintótica de las series "cuasi-integradas" y las integradas no difiere en sus implicaciones para muestras finitas o en otras palabras, las discontinuidades entre las raíces unitarias y las estacionarias (explosivas) sólo ocurre en el límite. Incluso desde el punto de vista de una valoración económica en términos de teoría de bienestar, la existencia de raíces unitarias y "cuasi-unitarias" es esencialmente equivalente. En un mundo con incertidumbre donde se descuenta el futuro, si la alternativa a raíces unitarias es una reversión a la media extremadamente larga, entonces la distinción entre la hipótesis nula de raíz unitaria y la alternativa no contendrá importantes consecuencias en términos de bienestar social. En este sentido, los enfoques que tratan de medir el "tamaño" de la raíz unitaria en la descomposición de una serie temporal en sus diversos componentes (véase p. ej. Cochrane (1988) y el trabajo del profesor Fernández Macho) me parecen muy fructíferos.

3. RAICES UNITARIAS Y CAMBIOS ESTRUCTURALES

Tal como señalan correctamente los profesores Escribano y Molinas, el supuesto simplificador adoptado en el artículo de que los procesos estocásticos tienen media cero, resulta excesivamente restrictivo y nada inócuo. Lo he adoptado para restringir la discusión a procesos con varianza explosiva, pero la existencia de componentes determinísticos (medias, tendencias, observaciones atípicas, etc.) puede desvirtuar algunos de los resultados del artículo. De nuevo en este frente, las noticias son buenas como lo muestra el número creciente de trabajos publicados sobre el tema (véase p. ej. Park y Phillips (1988) Perron (1989), Ouliaris et al. (1987) Banerjee et al. (1989) y el trabajo de los propios comentaristas mencionados).

En resumen, la teoría de la cointegración, a pesar del gran número de artículos escritos sobre el tema, se encuentra todavía en una etapa preliminar. El interés mostrado sobre el tema por personas del calibre de los comentaristas es sin duda una señal de su progresivo avance en un futuro próximo.

REFERENCIAS

ANDRES, J., MOLINAS, C. y D. TAGUAS (1990). "Una Función de Consumo Privado para la Economía Española: Aplicación del Análisis de Cointegración". Documento de Trabajo SGPE-D-90002. Dirección General de Planificación. Ministerio de Economía y Hacienda.

BANERJEE, A., LUMSDAINE, R. and J. STOCK (1989) "Recursive and Sequential Tests for a Unit Root: Theory and International Evidence" University of Florida (mimeo).

COCHRANE, J. (1988) "How Big in the Random Walk in GNP?" *Journal of Political Economy,* 96, 893-920.

DOLADO, J. and HIDALGO J.(1990) "The Asymptotic Distribution of the Iterated Gauss-Newton Estimators of an ARIMA Process" (mimeo) (de próxima aparición en *Econometric Theory* 6, no. 4).

HALL, A. (1989) "Testing for a Unit Root in the Presence of Moving Average Errors" *Biometrika* 76, 49-56.

HANSEN, B. (1990) "A Powerful, Simple Test for Cointegration Using Cochrane-Orcutt" Working Paper 230. University of Rochester.

OULIARIS, S., PARK, J. and P. PHILLIPS (1988) "Testing for a Unit Root in the Presence of a Mantained Trend" en B. Raj. (ed) *Advances in Econometrics and Modelling.* Needham: Kluwer Academic Press.

PARK, J. and P. PHILLIPS (1988) "Statistical Inference in Regression with Integrated Processes: Part 1 *Econometric Theory* 4, 468-497.

PERRON, P. (1989) "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Econometrica* 57, 1361-1402.

PHILLIPS, P. (1988) "Regression Theory for Near-Integrated Time Series" *Econometrica* 56, 1021-1044.

SAID, S. and D. DICKEY (1985) "Hypothesis Testing in ARIMA (p, 1, q) Models" *Journal of the American Statistical Association,* 80, 369-374.